

Bayesi talutav raskus: tõenäosuslik andmeanalüüs praktiku pilgu läbi

PÄRT PROMMIK^{1,2,3}, AARE MÄRTSON^{1,2}, HELGI KOLK^{1,2}, ÜLO MAIVÄLI⁴

¹Traumatoloogia ja ortopeedia kliinik, Kliinilise meditsiini instituut,
Tartu Ülikool

²Traumatoloogia ja ortopeedia kliinik, Tartu Ülikooli Kliinikum

³Sporditeaduste ja füsioteraapia instituut, Tartu Ülikool

⁴Tehnoloogiainstituut, Tartu Ülikool

Olulised mõisted

Tõenäosus – numbriline mõõt meie ebakindluse määrale parameet-riväärtuse kohta.

Eeljaotus (*prior distribution*) – tõenäosusjaotus, mida kasutatakse eelteadmise lisamiseks mudelisse.

Tõepärafunktsioon (*likelihood function*) – tõenäosusjaotus, mida kasutatakse andmete lisamiseks mudelisse. Mudeldab andmete sobivuse kõikvõimalike parameetriväärtustega.

Järeljaotus (*posterior distribution*) – tõenäosusjaotus, mis saadakse eeljaotuse ja tõepärafunktsiooni kombineerimisel.

Markovi ahelate Monte Carlo meetod (MCMC) – simulatsioonipõhi-ne meetod, millega võetakse juhuvalim järeljaotusest ja moodustatakse Markovi ahel.

Tõenäosusintervall (*credible interval*, TI) – Järeljaotusest arvatud vahemikhinnang, millesse mudeli arvates mahub etteantud tõenäosu-sega parameetri tegelik väärtus.

Sissejuhatus

Andmeanalüüsi taga, kui see toimub teaduses, on enamasti soov lasta andmetel heita valgust teaduslikele teooriatele. See protsess on kahe-suunaline, sest just teooriate valguses kogume ju neidsamu andmeid, ning kõik see toimub mudelite vahendusel [1]. Teadusliku mudeldamise edukus sõltub paljuski sellest, kui hästi kajastavad mudelid andmeid tootva protsessi erinevaid tahke. Kui meid huvitav looduslik protsess on

liiga keeruline, et seda ühe suure mudeliga haarata, siis võib seda hoo- mata lihtsamate mudelite ansambli abil, kus erinevad mudelid vastavad erinevatele küsimustele andmeid tootva protsessi erinevate tahkude kohta, ning aitavad hinnata ka üksteise robustsust. Samuti võimaldab selline lähenemine õppida halvasti käituvatelt mudelitelt, kuidas ehitada paremaid. Bayesiaanliku andmeanalüüsi õppimiseks soovitame kahte õpikut [2,3], millest esimene keskendub bayesiaanliku regressiooni praktikale ja teine selle sügavamale mõistmisele. Tartu Ülikooli kursus „Bayesiaanlik andmeanalüüs loodusteadlastele“ (LOTI.01.009) käsitleb samu teemasid. Senikaua kui andmetega kaasneb ebakindlus, sealhulgas mitte-redutseeritav looduslik varieeruvus, on meie mudelid stohhastilised ja ka nende väljund on paratamatult ebakindel. Bayesiaanlikus statistikas tähendab see, et hinnangud parameetriväärtustele tulevad meile tõenäosusjaotustena, ja et meile on oluline ebakindluse aus edasikandmine andmetest teaduslikesse järeldustesse. Seevastu sageduslikus statistikas arvutatakse parameetriväärtus punkthinnanguks ja ebakindlust kajastab enamasti sellele lisatud p - väärtus. Seega on traditsioonilise statistika fookus andmete redutseerimisel punktväärtusteks protseduuride abil, millel on teatud sageduslikud omadused, eelkõige hinnangute madal kallutatus ning kitsad usalduspiirid [4]. Vaatame kõigepealt lihtsat regressioonimudelit, mida kirjeldab võrrand $y = a + bx + \text{vealiige}$. Sageduslikus statistikas arvutame vähimruutude meetodil a ja b , kuid mitte vealiikme punktväärtused (vealiiget saab hinnata mudeli jääkidelt). Seda arvutust kutsutakse mudeli määramiseks (fittimiseks). Bayesiaanlikus kuues näeb see mudel välja nii:

$$y \sim \mathcal{N}(\mu, \sigma),$$

$$\mu = a + bx.$$

Esimene rida annab tõepäramudeli ehk ennustatava muutuja y jaotusmudeli, mis ütleb, et muutujale y omistatavad väärtused on määratud stohhastiliselt (tilde näitab stohhastilist seost) normaaljaotuse poolt, mille parameetrid on μ ja σ , millest esimene on identsussuhtes muutuja y aritmeetilise keskmisega ja teine vastab selle muutuja standardhälbele. Teine rida, nn protsessimudel, defineerib μ ümber muutuja y ja muutuja x deterministliku sõltuvussuhtena. Seega kaob

modelist μ ja selle asemele tulevad parameetrid a ja b . Seega ütleb täismudel, et muutuja y jaotus (ja seega ka tõepäramudel) on tingimuslik muutuja x väärtusest. Meie modelis määratakse kolm parameetrit (a , b ja σ) ja nõnda fikseeritud model on generatiivne, st sellest saab genereerida uusi muutuja y väärtusi, mis sõltuvad vastavalt protsessimudelile muutujale x omistatud väärtustest. Selline generatiivsus iseloomustab kõiki bayesiaanlikke mudeleid, kuid mitte kõiki sageduslikke analooge.

Bayesiaanlikus regressioonis vastab модели igale parameetrile eeljaotus – meie näites muutujatele a, b ja σ –, mille abil saame modelisse lisada andmetest sõltumatut informatsiooni võimalike parameetriväärtuste kohta. Näiteks tõusukoefitsient b eeljaotuse tsentreerimine nullile lisab modelile konservatiivsust, „tõmmates“ b järeljaotust nulli suunas seda rohkem, mida kitsam on see eeljaotus ja mida vähem me modelisse lisame valimiandmeid. Mõistlikult valitud eeljaotused vähendavad ohtu, et me valimivea või kallutatud andmete põhjal teeme halbu teaduslikke järeldusi. Kui meil on piisavalt kvaliteetseid andmeid ja piisavalt lihtne model, võime kasutada tasaseid (mitteinformatiivseid) või laiu (väheinformatiivseid) eeljaotusi, mis on valitud selliselt, et nad ei mõjutaks järeljaotuse kuju. Väheinformatiivsete eeljaotuste korral andmed justkui kõneleksid ise enda eest. Kolme parameetriga модели määramine MCMC ahelatega annab meile numbrilise juhuvalimi 3D järeljaotusest (Joonis 1A), kus andmetega ja eeljaotusega paremini kooskõlas olevad parameetriväärtuste kombinatsioonid on esindatud sagedamini kui halvema kooskõlaga kombinatsioonid. Seega on meil siin tegu kolmemõõtmelise järeljaotusega (Joonis 1B). Sellise mitmemõõtmelise järeljaotuse saab parameetri kaupa lahutada komponentideks, mida sageli näidatakse graafiliselt karvaste tõukudena, ehk meie näites 4 iseseisva MCMC ahelana, kus ahela liikumise iga murdepunkt vastab ühele parameetri väärtusele, mis on tõmmatud järeljaotusest (Joonis 1C). Joonisel liiguvad kõik neli ahelat iga parameetri juures kattuvatel trajektoridel ja „tõuk“ on ühtlaselt karvane, ilma siledete osadeta ja suurte väljalöökideta. Seega on meie ahelad jooksnud hästi, mistõttu koefitsientide hinnangud on algoritmiliselt usaldusväärsed. Järeljaotuse saab avaldada ka kolme eraldi tihedusfunktsioonina, kus iga jaotuse tipp

vastab selle parameetriväärtuse kõige tõenäolisemale väärtusele (Joonis 1D). Järeldaotustega (Joonis 1A), saab teha aritmeetilisi tehteid. Näiteks kui tahame teada, milline on muutuja y ennustatav keskvaartus juhul kui $x = 1$, liidame a ja b koefitsientide järeldaotuste vektorid $(a + b \cdot 1)$, mis annab meile järeldaotuse sellele y keskvaartusele.

Regressioonimodelit on lihtne laiendada (i) asendades tõepäramudelid normaaljaotuse mõne muu tõenäosusjaotusega (binoomjaotusega, lognormaaljaotusega, jms), (ii) modifitseerides protsessi mudelit, k.a mitte-lineaarseks, (iii) defineerides ümber rohkem kui ühe tõepäramudeli parameetri (näit μ ja σ), sama või erinevate protsessi mudelitega, seome muutujaga x varieeruvuse hinnangu, (iv) pannes mudelisse rohkem kui ühe ennustatava muutuja, (v) transformeerides ennustatava y -muutuja(d) näiteks logit-transformatsiooniga, (vi) lisades eeldaotustele meta-eeldaotusi muudame mudeli mitmetasemeliseks[2].

Töövahendid

Andmed ja uurimisküsimus

Bayeesanliku töövoogi illustreerimiseks kasutame näidet, kus uurisime reieluu puusaliigese piirkonna murruga haigete ravikäsitlemist Eestis [5,6]. Sellise murruga haigete prognoos on tõsine – taastumine võib võtta üle aasta ning suurel osal haigetest ei taastu varasem kehaline võimekus mitte kunagi, mistõttu väheneb või kaob nende iseseisvus igapäevatoimetustega hakkamasaamisel. Reieluu puusaliigese piirkonna murruga haigete ravi alussammas on järjepidev, koordineeritud taastusravi, milles meie uurimustöö tuvastas olulisi puudusi. Populatsioonipõhised retrospektiivsed andmed koguti Eesti Haigekassa andmebaasist. Uuringusse kaasati kõik 50 aastased ja vanemad isikud, kellel diagnoositi ajavahemikul jaanuar 2009 – september 2017 reieluu proksimaalse osa murru diagnoos. Diagnoos põhines Rahvusvahelise haiguste klassifikatsiooni 10. versiooni koodidel S72.0-2. Uuritavate luumurru diagnoosid valideeriti. Andmed sisaldasid nende haigete füsioteraapia kasutamist ravi kahes faasis, täpsemalt aktiiv- ja järeldravis. Ravi esimeses etapis (aktiivdravis) haige hospitaliseeritakse, opereeritakse ja seejärel alustatakse füsioteraapiaga. Võrdlemisi lühikesele aktiivdravile järgneb järeldravi, kus

peamine fookus on kehalise võimekuse taastamisel. Just järelravi ajal oli füsioteraapia haigetele halvasti kättesaadav. Meie uuringu põhieesmärk oli uurida, kas olukord paranes, ehk kuidas muutus füsioteraapia pakkumine üheksa-aastaselt perioodil 2009–2017.

Tarkvara

Bayesi teoreemi asemel määratakse mudeli koefitsiendid Markovi ahelate Monte Carlo (Markov chain Monte Carlo, MCMC) simulatsioonide abil, mis korjab juhuvalimi otse järeljaotusest. MCMC ahelate kasutamine on muudetud nii lihtsaks, et enamasti pole kasutajal vaja sellele palju mõelda. Siin omasid suurt rolli esimesed lihtsalt kasutatavad Markovi ahelaid jooksutavad programmid Bugs, winBugs ja JAGS, mille on praeguseks suuresti asendanud Stan, mis kasutab Hamiltonian Monte Carlo meetodit [7]. Kuna programmi Stan on keerulisevõitu õppida, kasutatakse mudelite ehitamiseks sageli abiprogramme, millest levinumad on R::rstanarm ja R::brms, mis võimaldavad mudelid kodeerida R::lme4 ja sarnaste pakettide süntaksit kasutades [8–10]. Kogu protseduur – mudeli programmeerimine, jooksutamine ja töö mudeli objektiga – toimub kasutaja jaoks Stani programmi avamata (see töötab vaikselt nn kapoti all, võttes sisse abiprogrammis spetsifitseeritud mudeli ja väljastades abiprogrammi järeljaotuse ja mudeli kvaliteedinäitajad). R::rstanarm keskendub levinumate sageduslike mudelite analoogidele (k.a ANOVA ja elulemuse mudelid) [10,11]. Brms annab laiemad valikud mudelite ehitamiseks ja on ka laiemas kasutuses.

Bayesiaanlik töövoog

Esmane andmeanalüüs

Esmane ehk eksploraatiivne andmeanalüüs annab aimu nii andmete kvaliteedi, kui formaalse mudeldamise võimalike tulemuste kohta. See omakorda võimaldab paremini ära tunda mudeldamisel juhtuda võivaid vigu, olgu need siis algoritmilist või tõlgenduslikku laadi. Lisaks aitab parem arusaam andmetest meil oma mudeleid paremini spetsifitseerida, näiteks suunata valikuid ühe- ja mitmetasemeliste

modelite vahel [12]. Me analüüsisime 11 461 reieluu puusaliigese piirkonna murruga haige taastusravi. Selgus, et kui aktiivravi ajal ei saanud 24% haigetest füsioteraapiat, siis järelravi käigus ei saanud tervelt 60% haigetest füsioteraapiat ning ülejäänud said seda võrdlemisi väikeses koguses (mediaan 6 tundi, 25- ja 75-protsentiil: 3, 11). Järelravi ajal füsioteraapiat mittesaanute osakaal vähenes üheksa-aastase perioodi jooksul ja seda > 10 tunni saanute osakaal suurenes (Joonis 2A). Lisaks erinevad maakonnad tugevalt taastusravis ja selle trendides [6], mis viitab sellele, et mitmetasemelised mudelid võiksid hästi sobida järelravi ajal saadud füsioteraapia muutuste analüüsiks. Mudeldamise seisukohast on oluline ka peamise tulemuslikkuse mõõdiku – järelravi ajal saadud füsioteraapiatundide arvu – jaotus. Jooniselt 2B näeme, et see sisaldab palju nulle, näidates mitmeosalise regressioonanalüüsi vajalikkust.

Tõepärafunktsiooni leidmine

Bayesiaanlikus regressioonis ehitame oma mudelid – olgu need lineaarsed või mittelineaarsed, ühe- või mitmetasemelised – samale mõisteli-sele alusele ja jooksutame neid sama MCMC algoritmiga. Esmalt spetsi-fitseerime andmemudeli ehk tõepärafunktsiooni. Tõepärafunktsioon mudeldab sõltuva tunnuse (y) väärtusi tingimuslikena x -muutujate (prediktorite) väärtustest. Kuna meil on prediktoriks nullide all lookas ja tugevalt paremale kiivas pidev tunnus (Joonis 2B), siis analüüsi-me ajalisi trende kolmes osas: (i) füsioteraapiat saanute osakaal, (ii) saadud füsioteraapiatundide arv ja (iii) saadud füsioteraapiatundide varieeru-vus patsientide vahel. Niisiis proovime leida kõigile analüüsi osadele sobivad tõepärafunktsioonid. Brms pakub mudeldamiseks 40 erinevat tõepäramudelit ning kasutajal on võimalik ka oma soovi järgi spetsifi-tseerida mudeli jaotuslik vorm. Kuna kõik bayesaanlikud mudelid on generatiivsed (mudelist saab genereerida uusi andmepunkte), aitab sobivat tõepärafunktsiooni valida määratud mudelist *in silico* genereeri-tud andmete võrdlemine valimiandmetega (posterior predictive check). Kui võrdlus näitab halba kokkulangemist, siis võime muuta tõepära-funktsiooni, sõltumatuid tunnuseid, nendevahelisi seoseid või lisada mudelisse konditsionaalseid varieeruvusi. Sõltuvalt analüüsi eesmär-

gist, võivad meid huvitada andmejaotuste sobivuse erinevad aspektid. Näiteks kui oleme huvitatud keskväärtustest, siis võrdleme in silico valimite keskväärtusi empiirilise valimi keskväärtusega, kui meid huvitavad varieeruvused või näiteks maksimumväärtused, siis võrdleme hoopis neid. Mõneti käib tõepärafunktsiooni valimine katse-eksitus meetodil. Esimese analüüsi osana huvitas meid, kuidas füsioteraapiat saanute osakaal muutus järelravi ajal. Selleks muutsime sõltuva tunnuse binaarseks („Ei saanud“, „Sai“) ja proovisime teda mudeldada Bernoulli jaotusega. See mudel näeb brms-i koodis välja nii:

```
brm(teraapia_jarelravi_binaarne ~ aasta,
family = „bernoulli“, link = „logit“, data = andmed)
```

Analüüsi teine osa keskendus järelravi ajal saadud füsioteraapia mahu hindamisele. Mudelile sobiva jaotusliku vormi leidmiseks proovisime kolme tõepärafunktsiooni: normaaljaotus, kallutatud normaaljaotus ja lognormaaljaotus. Lisaks täiendame kõiki kolme mudeleid selliselt, et nad üheaegselt mudeldaksid sõltuva tunnuse varieeruvuse sõltuvust sõltumatutest tunnusest. Sellised mudelid annavad üheaegselt hinnangu sõltuvalt tunnusele ja ka selle varieeruvusele igal sõltumatu tunnuse väärtusel. Vastavad mudelid spetsifitseeritakse brms-is järgmiselt:

```
brm(bf(teraapia_jarelravi ~ aasta, sigma ~ aasta),
family = „gaussian“, data = andmed)
brm(bf(teraapia_jarelravi ~ aasta, sigma ~ aasta),
family = „skewnormal“, data = andmed)
brm(bf(teraapia_jarelravi ~ aasta, sigma ~ aasta),
family = „lognormal“, data = andmed)
```

Joonisel 3 näeme ülaloodud mudelite ennustuse sobivusi originaalandmetega. Neist kolmest näitab selgelt parimat parimat sobivust lognormaaljaotusel põhinev mudel, mis taastab võrdlemisi täpselt nii mediaan-füsioteraapiatunnid kui ka nende varieeruvuse (Joonis 3C).

Lõpetuseks, brms lubab meil mudelite ennustusi reparametriseerida [13], mis võimaldab valida, millises skaalas me soovime mudelist ennustusi saada. Näiteks lognormaalse mudeli füsioteraapiatundide ennustusi saab näidata nii keskmisena kui ka mediaanina, varieeruvust saab näidata standardhälbena, variatsioonikoefitsendina ja geomeetrilise standardhälbena [14]. Käesolevas analüüsis reparametriseerime

lognormaalse mudeli selliselt, et see annaks tulemused loomulikus skaalas ehk tundides: füsioteraapia tundide mahu mediaantundidena ja varieeruvuse standardhällbena .

Protsessimudelite ehitamine

Ajaliste trendide mudelite spetsifitseerimisel arvestasime kolme aspektiga. (i) Analüüsis tuleb arvestada aktiivravi ajal saadud füsioteraapiaga, mis võis 9-aastaselt perioodil samuti muutuda (haige taastumise seisukohast ei ole suurt vahet, millises ravifaasis tema taastusravi läbi viiakse). (ii) Andmestruktuur oli hierarhiline: haiged said taastusravi erinevates maakondades, millistes pakutakse taastusravi suhteliselt iseseisvalt. (iii) Meie varasem analüüs näitas, et nende haigete taastusravi ja selle ajalise muutused erinesid maakonniti. Seega on haige elukoht oluline. Eelnevast tulenevalt jooksutasime kuus erineva keerukusega Bernoulli (ravi saajad) ja lognormaalsel jaotusel (teraapiatundide arv) põhinevat mudelit, millest viimane sisaldas ainult positiivseid väärtusi. Mudelid spetsifitseeriti järgnevalt:

$$teraapia_{jarelravi} \sim 1$$

$$teraapia_{jarelravi} \sim aasta$$

$$teraapia_{jarelravi} \sim aasta + teraapia_{aktiivravi}$$

$$teraapia_{jarelravi} \sim aasta + teraapia_{aktiivravi} + maakond$$

$$teraapia_{jarelravi} \sim aasta + teraapia_{aktiivravi} + (1|maakond)$$

$$teraapia_{jarelravi} \sim aasta + teraapia_{aktiivravi} + (aasta|maakond)$$

Lognormaalsed mudelid mudeldasid paralleelselt ka standardhälvet, kasutades samu sõltumatuid tunnuseid.¹⁹

Kuna kõik mudelid jooksid probleemideta, jooksutasime neid brmsi vaike-eeljaotustega. Regressiooni puhul tuleb alati veenduda mudeldamise õnnestumises, sest veateadetega mudelite ennustused ei pruugi olla usaldusväärsed. Esiteks tuleb tähelepanu pöörata veateadetele, mis viitavad MCMC ahelate mitte-konvergeerumisele või ohule, et ahelad ei ole külastanud kõiki multi-dimensionaalse posteeriori osi. Veateated võivad tähendada mudeli konstrueerimise katastroofilist ebaõnnestumist, aga ka olla pelgalt leebed hoiatused. Enamasti viitavad tõsisemad probleemid mudeli sobitamisel andmetega sellele, et mudel on valesti spetsifitseeritud ja annavad seega alust mõne hooletusvea parandamiseks. Erinevatest veateadetest annavad ülevaate [3] ja nende lahendamiseks on Stani arendusmeeskond välja andnud soovitusi [15]. Mõnikord lahenevad mudeli konstrueerimisega seotud probleemid andmete transformeerimisega (näiteks y -tunnuse tsentreerimisel, standardiseerimisel, logaritmimisel või logit-skaalasse transformeerimisel) [16]. Teinekord tuleb muuta eeljaotus informatiivsemaks ehk rangemalt piirata mudeli erinevate parameetrite ulatusi või protsessimudel reparametriseerida [2].

Mudelite võrdlemine

Parima mudeli valik sõltub analüüsi eesmärgist. Põhjusliku analüüsi puhul kasutame selleks suunatud atsüklilisi diagramme (*directed acyclic graph* ehk DAG), mis aitavad meil algoritmiliselt otsustada, milliseid prediktoreid mudelisse panna. Parim viis prediktorite vaheliste suhete spetsifitseerimiseks (näiteks haige elukohamaakonna tunnuse mudeldamine eraldi tasemena või erinevate prediktorite vaheliste interaktsioonide lisamine) otsustatakse sobitades mudeli struktuuri meie täpsema arusaamaga andmeid genereeriva protsessi struktuurist.

¹⁹ `brm(bf(teraapia_jarelravi ~ aasta + teraapia_aktiivravi +
(aasta | maakond),
sigma ~ aasta + teraapia_aktiivravi + (aasta | maakond)), family =
,lognormal')`

Tabelis 1 toome kuue mudeli bayesaanliku *leave-one-out* (LOO) ristvalideerimise võrdluse. LOO annab hinnangu mudeli valimivälise sobivuse kohta andmetega. Tabelit 1 interpreteerime nii: kui se_{diff} on vähemalt 3 korda suurem kui $elpd_{diff}$, siis mudel on suure tõenäosusega eelmisest oluliselt halvem. $Elpd_{diff} = 0$ tähistab võrreldavate mudelite seast „parimat“, kuigi suhteliselt parim mudel ei pruugi olla „hea“ mudel. Seega näitab tabel 1, et Bernoulli mudelite puhul on kolm esimest võrdsed ja ülejäänud kolm on neist palju halvemad ja omavahel sarnased. Mõlemad mudelid (Bernoulli ja lognormaalsed) järjestuvad loo järgi samamoodi. Samamoodi saame hinnata standardhälbe (σ) mudeldamise mõju mudeli headusele. Näiteks saame Tabel 1 parima lognormaalse mudeli ehitada selliselt, et see mudeldab sõltuvalt x -prediktoritest y -muutuja keskväärtust aga mitte sigmat. Sigma mittemudeldamine halvendas tunduvalt mudeli sobivust ($elpd_{diff} = -193.7$ ja $se_{diff} = 19.1$), mida võib tõlgendada kinnitusena, et meie arusaam andmeid tootva protsessi hierarhisest struktuurist on adekvaatne.

Tabel 1. Mudelite võrdlemine bayeesanliku *leave-one-out* ristvalideerimisega

Mudel	Bernoulli		Lognormaalne	
	$elpd_{diff}$	se_{diff}	$elpd_{diff}$	se_{diff}
$\sim aasta + teraapia_{aktiivravi} + (aasta maakond)$	0,0	0,0	0,0	0,0
$\sim aasta + teraapia_{aktiivravi} + (1 maakond)$	0,0	0,0	-32,9	8,9
$\sim aasta + teraapia_{aktiivravi} + maakond$	-0,1	0,5	-33,3	9,2
$\sim aasta + teraapia_{aktiivravi}$	-492,6	30,9	-308,7	25,7
$\sim aasta$	-521,9	31,7	-342,7	27,3
~ 1	-611,4	34,1	-376,1	28,3

Samu sõltumaid tunnuseid kasutati standardhälbe paralleelseks mudeldamiseks lognormaalses mudelites; $elpd$ näitab määratud mudeli sobivust andmetega, ehk mudeli ennustuste täpsust; $elpd_{diff}$ – oodatud logaritmitud ennustatavate tiheduste erinevus (expected log-predictive density); se_{diff} annab standardhälbe $elpd_{diff}$ suuruse hinnangu järelejaotusele;

Järeljaotuste informatiivsus

Bayesaanliku regressioonimudeli parameetrid antakse järeljaotuse kujul, mis sisaldavad kogu mudelis sisalduvat teavet nende parameetrite kohta. Järeljaotuse kasutamine annab meile paindlikud ja lihtsalt teostatavad võimalused analüüsi tulemuste raporteerimiseks koos tõenäosusintervallidega. Näiteks arvutasime y -tunnuse ennustuse vaadeldud perioodil algus- ja lõppaasta kohta (Joonis 4A), ning nende kahe jaotuse pealt arvutame mitmesuguseid statistikuid, eesmärgiga kirjeldada taastusravi ajaliste trende. Me saame järeljaotust kirjeldada kasutades tavapäraseid keskmist iseloomustavaid statistikuid (mood, mediaan, aritmeetiline keskmine), loomulikult koos tõenäosusintervalliga (TI), mille sageduslik analoog on usaldusintervall. Näiteks 2009. aastal said haiged järelravi ajal mediaanina 4,4 tundi [95% TI: 3,2; 5,9] füsioteraapiat ja 2017. aastal 5,3 tundi [3,6; 7,6] (Joonis 4A). Teiseks saame olemasolevatest järeljaotustest arvutada uusi huvitavaid järeljaotusi. Näiteks taastusravi 9-aastase trendi kirjeldamiseks lahutame ühest järeljaotusest teise, saades nõnda efekti suuruse

$$ES = Y^{2017} - Y^{2009}$$

järeljaotuse (Joonis 5B, $ES = 1.0$ [95% TI: -0.6; 2.7]). Kõigi mudelite 9-aasta ES-d on näidatud Joonise 4 paneelidel C–E. Järeljaotusi saab kasutada ka spetsiifiliste hüpoteeside testimiseks. Kasutame selleks vabade tõusudega hierarhilist mudelit. Näiteks võime testitava hüpoteesina defineerida kui kliiniliselt ebaolulise muutuse trendi, mis on väiksem kui 2 tundi 9 aasta vältel. Kuna ES järeljaotuse tõenäosusmassist jääb 93% selle hüpoteesi poolt defineeritud vahemikku, siis on kliiniliselt olulise positiivse ajalise muutuse tõenäosus 0,07 (Joonis 4F). Arvutame ka bayesaanliku p -väärtuse, ehk testime hüpoteesi, et keskmiste ravitundide arv on hoopis ajas langev. Selle hüpoteesi tõenäosus on 0,11 (Joonis 4G). Samas, tõenäosus, et trend on ajas langev ja samas kliinilises mõttes olulise suurusega, on kõigest 0,0005 (Joonis 4F). Seega, ehkki meie testid viitavad kokkuvõttes sellele, et mediaan-füsioteraapiatunnid pigem ei kasvanud 9 aasta jooksul kliiniliselt olulisel määral, jätavad nad siiski ka arvestatavat tõenduslikku kaalu hüpoteesile, mille kohaselt selline kasv ikkagi toimus.

Mudeli koefitsiendid

Eelmises peatükis nägime, et mudelist järelduste tegemine hõlmab eeskätt tööd mudeli ennustustega, mitte koefitsientide otse tõlgendamist. Siiski tasub, eriti mitmetasemeliste mudelite puhul, ka koefitsientidele pilk peale heita. Siin toome loo kriteeriumi järgi meie parima mudeli

$$bf(teraapia_{tunnid} \sim aasta + teraapia_{aktiivravi} + (aasta|maakond),$$

$$sigma \sim aasta + teraapia_{aktiivravi} + (aasta|maakond))$$

koefitsiendid logaritmilises skaalas koos 95% tõenäosusintervallidega (Tabel 2). See mudel jaotab patsientide üldpopulatsiooni ja maakonna taseme vahel nii varieeruvust kui varieeruvuse varieeruvust. Varieeruvust jaotades võrdleme maakondade keskmiste ravitundide varieeruvust sellest üle jääva varieeruvusega, mis iseloomustab maakonnast sõltumatut indiviidide vahelist varieeruvust. Näiteks näeme, et populatsiooni tasemel intercepti varieeruvus $sigma_{Intercept}$ on ca 5 korda suurem kui maakonna tasemel $sd(Intercept)$ ehk vabaliikmete standardhälve, mis saadakse 15 maakonna keskmiste ravitundide varieeruvusest. Seega on mudeli arvates ca $\frac{1}{6}$ ravitundide varieeruvusest seletatav maakonna mõjuga ravitundidele ja ülejäänud $\frac{5}{6}$ jääb populatsiooni tasemele seletust ootama. Seevastu varieeruvuse tõus ajas jaguneb üsna võrdselt populatsiooni taseme $sd(sigma_{aastatsentreeritud})$ ja maakonna taseme $sd(sigma_{aastatsentreeritud})$ vahel, mis viitab sellele, et maakonna tasemel toimuvad muutused omavad suuremat rolli seletamiseks ebavõrdsuse kasvu tõusu, kui maakonna tasemel mõjud suudavad seletada ebavõrdsust ennast. Lisaks näitab maakonna tasemel koefitsient $sd(\sigma_{Intercept})$ varieeruvuse varieeruvust maakondade vahel, mis osutub olema ligikaudu võrdne keskmise ravitundide arvu varieeruvusega maakondade vahel ($sd(Intercept)$). See varieeruvuse varieeruvus käib selle kohta, millisel määral erineb ravitundide kasutamise varieeruvus maakondade vahel, ehk kui palju erineb ravitundide jaotamise ebavõrdsus maakondade vahel.

Ja lõpuks, korrelatsioon $cor(Intercept, aasta_{tsentreeritud})$ näitab, et mudeli hinnang võimalikule korrelatsioonile keskmisele ravitundide arvule maakonnas ja selle muutusele ajas on liiga ebatäpne, et

seda mõistlikul viisil tõlgendada. Sama järelduse oleme sunnitud tegema korrelatsiooni kohta ravitundide varieeruvuse ja selle kasvu vahel maakondade lõikes. Kuigi, meil on mudeli arvates väga nõrk tõendusmaterjal, et esimene korrelatsioon võiks olla positiivne (maakondades, kus antakse rohkem ravi inimese kohta, see näitaja pigem tõuseb ajas) ja et teisel juhul võiks korrelatsioon olla negatiivne (maakondades, kus esineb suurem ebavõrdsus ravi jagamisel inimeste vahel, see ebavõrdsus pigem langeb ajas).

Tabel 2. Mudeli

$$bf(teraapia_{tunnid} \sim aasta + teraapia_{aktiivravi} + (aasta|maakond),$$

$$sigma \sim aasta + teraapia_{aktiivravi} + (aasta|maakond))$$

koefitsiendid (maakonna tasemel intercepte ja tõuse ei näidata).

Koefitsient	Punkthinnang	Alumine TI	Ülemine TI
Populatsiooni tasemel koefitsiendid			
<i>Intercept</i>	1,938	1,751	2,126
<i>sigma</i> _{Intercept}	1,948	1,724	2,175
<i>aasta</i> _{tsentreeritud}	0,042	0,013	0,067
<i>teraapia</i> _{aktiivravitsentreeritud}	0,043	0,029	0,059
<i>sigma</i> _{aastatsentreeritud}	0,079	0,035	0,121
<i>sigma</i> _{teraapiaaktiivravitsentreeritud}	0,043	0,020	0,070
Maakonna tasemel koefitsiendid			
<i>sd</i> (<i>Intercept</i>)	0,340	0,223	0,532
<i>sd</i> (<i>sigma</i> _{Intercept})	0,407	0,270	0,627
<i>sd</i> (<i>aasta</i> _{tsentreeritud})	0,041	0,022	0,073
<i>sd</i> (<i>sigma</i> _{aastatsentreeritud})	0,066	0,033	0,116
<i>cor</i> (<i>Intercept</i> , <i>aasta</i> _{tsentreeritud})	0,178	-0,417	0,674
<i>cor</i> (<i>sigma</i> _{Intercept} , <i>sigma</i> _{aastatsentreeritud})	-0,351	-0,814	0,284

Eeljaotuse mõju mudelile

Bayesaanlike mudelite määramine sõltub mitte ainult andmetest (tõepärafunktsiooni kaudu), vaid ka eeljaotusest. Range ehk kitsas eeljaotus võib mudeli parameetri järelejaotust olulisel määral nihutada ja/või kitsendada. Sellegipoolest kasutatakse rangeid eeljaotusi, mis lisavad mudelisse olulist valimiandmetest sõltumatut teaduslikku informatsiooni, vähe. Sellel on kolm võimalikku põhjust: (i) meil lihtsalt pole teaduslikku teadmist, mida mudelisse panna, (ii) me

ei oska oma mitte-matemaatilist teadmist eeljaotuse matemaatilisse vormi valada ja (iii) meil on nii palju kvaliteetseid andmeid, et eeljaotusse pandud lisainfo ei mõjutaks niikuinii järeljaotuse kuju. Sellistel juhtudel on populaarsed nõrgalt informatiivsed eeljaotused, mis küll ei kanna palju teaduslikku teadmist, kuid soodustavad mudeli parameetrite kitsendamise selle algoritmilist määramist MCMC meetoditega – eriti väheste andmete korral [16,17]. Kuna eeljaotus võib eelteadmise peegeldajana meid viia valedele radadele, tuleks kontrollida selle mõju järeljaotusele. Selline kontroll näitab ka andmete mõju ulatust määratud mudelile, mida on eriti keerulisemate mudelite korral hea teada. Me kasutasime brms-i eeljaotuste vaikeäätteid, sest meil on suur ja väga kvaliteetne andmestik. Range eeljaotuse mõju demonstreerimiseks võrdleme kahte versiooni sama protsessimudeliga hierarhilisest Bernoulli mudelist

```
brm(teraapia_jarelravi_binaarne ~
aasta + teraapia_aktiivravi + (aasta | maakond), family =
„bernoulli“, link = „logit“, data = andmed).
```

Ühe neist määrame brms-i vaike-eeljaotusega ja teisele anname range ja ebamõistliku eeljaotuse, mis illustab Eesti haigete ravikäsitlust. See ebasiiras eeljaotus ütleb, et me usume tõenäosusega 0,96, et Eestis taastusravi saajate osakaal on $> 75\%$ ja et me usume vaid 0,0008 tõenäosusega, et see on $< 50\%$ (vastavalt taastusravi ravijuhistele peaks see olema lähedal 100%-le). Joonis 5 annab ülevaate kahe mudeli eeljaotustest, ennustustest ja ennustuste erinevust. Näeme, et meie suur andmestik lahustas pea täielikult eeljaotuse mõju: range eeljaotusega mudelis tõusis teraapia saajate osakaal keskmiselt vaid 3,6 protsendipunkti võrra [95% TI: – 11, 1; 19, 1] (Joonis 5G). Joonis 5A näitab, et brms-i vaike-eeljaotus vabaliikmele (see näitab ravi saajate fraktsiooni), mis on mudelis defineeritud logit-skaalas, on tõenäosuste skaalasse transformeeritult nõrgalt informatiivne, eelistades ekstreemseid tõenäosusi. Sellest eeljaotusest ei sõltu meie andmete korral kuigi palju, aga peame arvestama, et väikeste andmestike puhul võib see tõmmata hinnanguid märgatavalt null- või ühiktõenäosuse suunas. Igal juhul, mida vähem andmeid, seda suuremat tähelepanu tuleb osutada informatiivsete eeljaotuste

ehitamisele.

Teaduslike järelduste tegemine

Rakendusliku statistika üks peamisi eesmäärke on mudeliennustuste põhjal teaduslike järelduste tegemine. Sageli on siin tegu nn faktivastaste (*counterfactual*) ennustustega, kus ennustatakse võrdlevalt täpselt samade näitajatega näidis-indiviidi väljundit, kui ta kuuluks katsegruppi ja kui ta kuuluks kontrollgruppi. Just sellised ennustused võimaldavad põhjuslikke tõlgendusi ehk kas katsetöötlus mõjutaks seda kujutletavat indiviidi põhjuslikult [18]. Kokkuvõttes tähendab see, et me teeme järeldusi suure maailma kohta, kasutades selleks nii mudel-ennustusi kui teadmisi, mida me mudelisse ei pannud. Eksploratiivne andmeanalüüs andis aimduse, milliseid trende võiksid regressioonimudelid edasises analüüsis näidata. Mudelitest *in silico* genereeritud andmete võrdlemine sõltuva tunnuse tegeliku jaotusega aitas meil valida tööpärafunktsioonid ja mudelite deterministliku osa struktuurid, mis võimaldavad realistlikke mudelipõhiseid ennustusi. Mitme mudeli tulemuste üheskoos vaatlemine andis parema arusaama tegeliku mõju suunast ja suuruselt. Eeljaotuse mõju hindamine kinnitas meie tulemuste valiidsust, näidates, et andmete suur hulk lahjendas ka väga tugeva eeljaotuse mõju sisuliselt olematuks. Mudelite ansambli põhised konsistentsed ajalised trendid võimaldavad teha järgmisi teaduslikke järeldusi:

1) füsioteraapia saajate osakaal suurenes järelravi ajal 28 protsendipunkti võrra [95% TI:16;39], mis annab tunnistust riiklikul tasemel toimunud pingutusest parema järelravi tagamiseks.

2) Kuid füsioteraapiat saanute teraapiatundide mediaan ei kasvanud kuigivõrd ning

3) füsioteraapiatundide jagamine patsientide vahel muutus üheksa-aastaselt perioodil 1,9 korda ebaühtlasemaks [95% TI:1,3;2,6] ja patsiendi tasemel saadud teraapiatundide standardhälve suurenes 4,6 tunni võrra [95% TI:2,0;7,0], mis viitab kohalike ravikeskuste tasemel aset leidva ebaõigluse süvenemisele raviressursi jagamisel patsientide vahel.

4) Mitmetasemelisest mudelitest saadud maakondlikud hinnangud meid huvitavatele parameetritele näitavad suurt varieeruvust maakon-

dade vahel kõigis kolmes näitajas ehk seda, et nii lisandunud raviressurss kui selle kasutamise efektiivsus liikus kohalikule tasemele väga ebaühtlaselt (vt [6] joonis 4 ja lisajoonis 5). Üldiselt on Lõuna-Eesti tunduvalt paremas seisus kui rannikuga piirnevad maakonnad, eriti Harjumaa, kus ravi saamise ebavõrdsus on väga suur ja ravi mittesaanute osakaal on väga kõrge.

Diskussioon

Käesoleva töö eesmärk oli anda kokkuvõtlik ülevaade bayesaanliku andmeanalüüsi töövoost, kasutades tegeliku taastusravi andmeid. Pole kahtlust, et bayesaanliku analüüsi töövoog on töömahukam tavapärasest. Teisalt aitab selline mitme-etapiline ja mudelite ansambli tõlgendamisel põhinev lähenemine ehitada tugevama vundamendi teaduslikele järeldustele. Kuigi klassikaliste meetoditega jõutakse sageli sarnaste teaduslike järeldusteni, on bayesaanlikul andmeanalüüsil mõningaid eeliseid [19]: (i) suurem paindlikus mudelite ehitamisel võimaldab kodeerida mudeleid, mis paremini kajastavad andmeid genereerivaid protsesse; (ii) mudelite määramise suurem tehniline edukus on eriti oluline keerulisemate mitmetasemeliste mudelite korral, eriti neil, mida arvutatakse vähestel andmetel; (iii) määratud mudelist saab genereerida *in silico* andmeid ja võrrelda neid valimiandmetega, mis lihtsustab mudelistruktuuri optimeerimist; (iv) mudelite järeljaotused on informatiivsemad kui punkthinnangud usaldusintervallidega või *p*-väärtustega, ning need võimaldavad vastata laiemale valikule teaduslikele küsimustele; (v) bayesaanlik traditsioon soosib mudelite võrdlemist, mis meeldib teadlastele, kellele on kombeks võrrelda erinevate hüpoteeside sobivust andmetega; (vi) teaduslike teadmiste lisamine mudelisse eeljaotuse kujul aitab vähendada järelduste ebakindlust ja/või kallutatust; (vii) bayesiaanlik analüüs, mis ei tõmba ületamatut piiri andmete ja mudeli parameetrite vahele, annab hea viisi puuduvate andmete imputeerimiseks (neid mudeli parameetriteks ümber defineerides), aga ka meta-analüüsiks; (viii) Bayesiaanlik tõenäosus, mis käib otse parameetriväärtuste kohta, on teadlase jaoks loomulikum ja lihtsamini kasutatav kui sageduslik tõenäosus, mis käib andmete kohta null-hüpoteesi kehtimise korral

[20]. Bayesaanlikul lähenemisel on leitud olema ka puudusi, sealhulgas sügavalt teoreetilisi [21,22]. Meie arvates on praktilist laadi puudused, mis reaalselt mõjutavad teadlast, kes rakendab bayesiaanlikku töövoogu, järgmised:

1) Puudub sisse töötatud vorm bayesiaanlike tulemuste avaldamiseks. Kuigi leidub publikatsioone, mis pakuvad reegleid selleks, lonkab siiani nende ühtne rakendamine praktikas [23,24]. Siin aitab analüüsikoodi täismahus avaldamine artikli lisana, soovitavalt koos lühikese õpetusega, kuidas seda koodi kasutada ja selles olevaid mudeliobjekte avada (vt näiteks [25]).

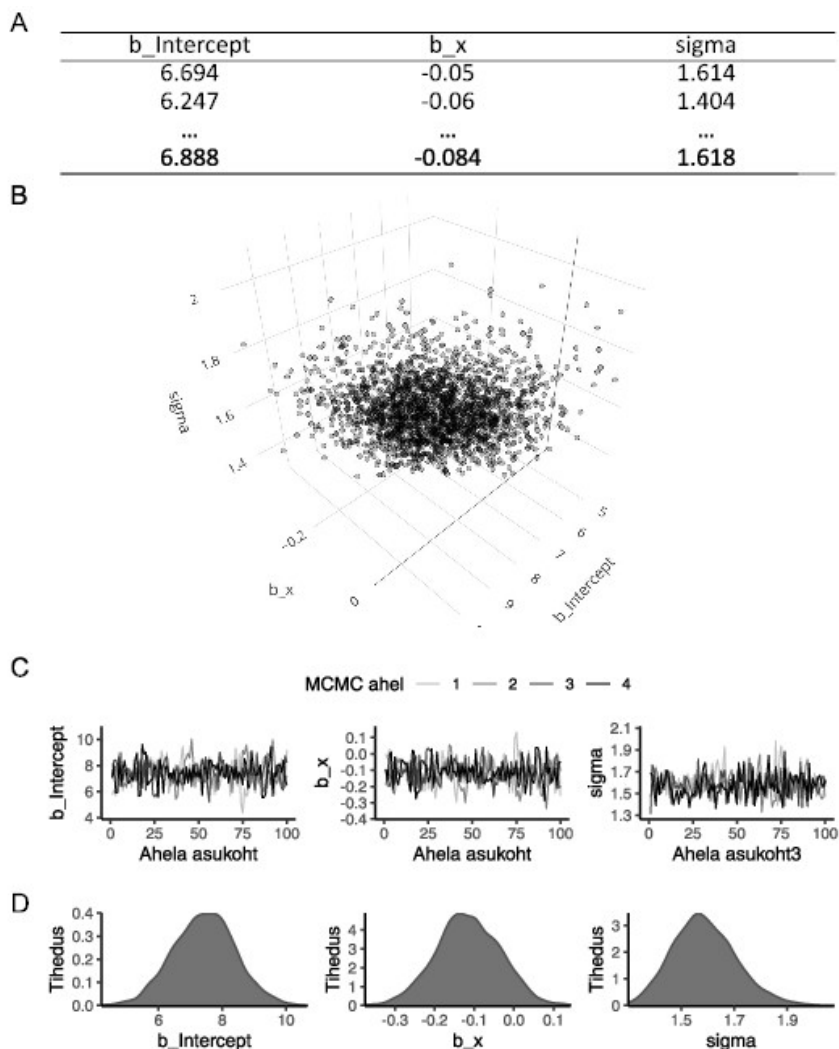
2) Informatiivsete eeljaotuste defineerimine keerukamatele mudelitele võib osutuda keerukaks (vt joonist 6A). Selle probleemiga tegeletakse aktiivselt.

3) Töö mudelite ansambliga loob ohu mudelite ülemääramiseks ja seega ületõlgendamiseks. Kui me eelistame tõlgendada mudeleid, mis paremini sobituvad meie valimiandmetega, siis jääb alati oht, et meie valimiandmed valetavad ja me teeme liiga julgeid järeldusi nende sobivuse põhjal mudelitega, millised me oleme välja valinud just sellesama sobivuse tõttu. Paraku alternatiiv, kus me määrame ühe mudeli ja elame siis selle mudeli ennustustega hoolimata sellest, kui hästi see mudel generatiivselt töötab või andmete struktuuriga sobitub, ei tundu meile kuigi ahvatlev. Mudelite ülemääramist aitab vältida eelkõige mudeli sobivus andmeid genereeriva protsessiga, eriti mitmetasemeliste mudelite korral. Sellega seoses on just bayesiaanlikus statistikas eriti oluline regressiooni sidumine formaalsete põhjuslike mudelitega [26].

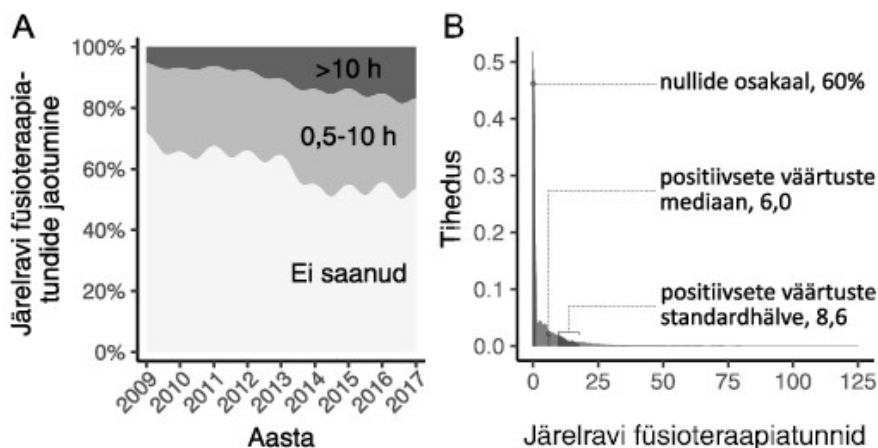
4) Ajakirjade retsensendid kipuvad võtma bayesiaanlikku andmeanalüüsi kriitikavabalt, mis langetab kokkuvõttes avaldatud tööde kvaliteeti või, mis veel hullem, tegema bayesiaanlikele töödele sagedaslikku kriitikat, mis raskendab avaldamist, lisamata midagi kvaliteedile.

5) Bayesiaanlikud mudelid jooksevad aeglaselt. Meie keerulisemate mudelite läbiarvutamine võtab sülearvutil kuni paar tundi. Lihtsad mudelid jooksevad tavaliselt sekundite kuni minutite jooksul.

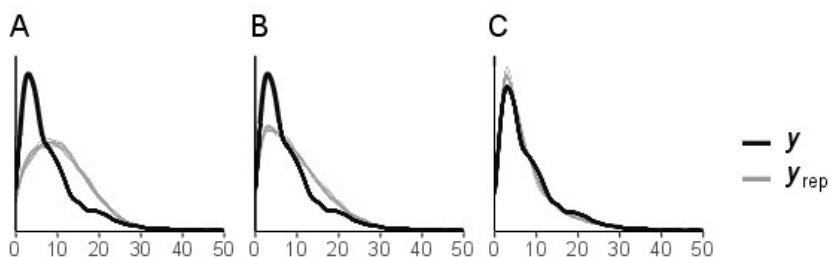
Joonised



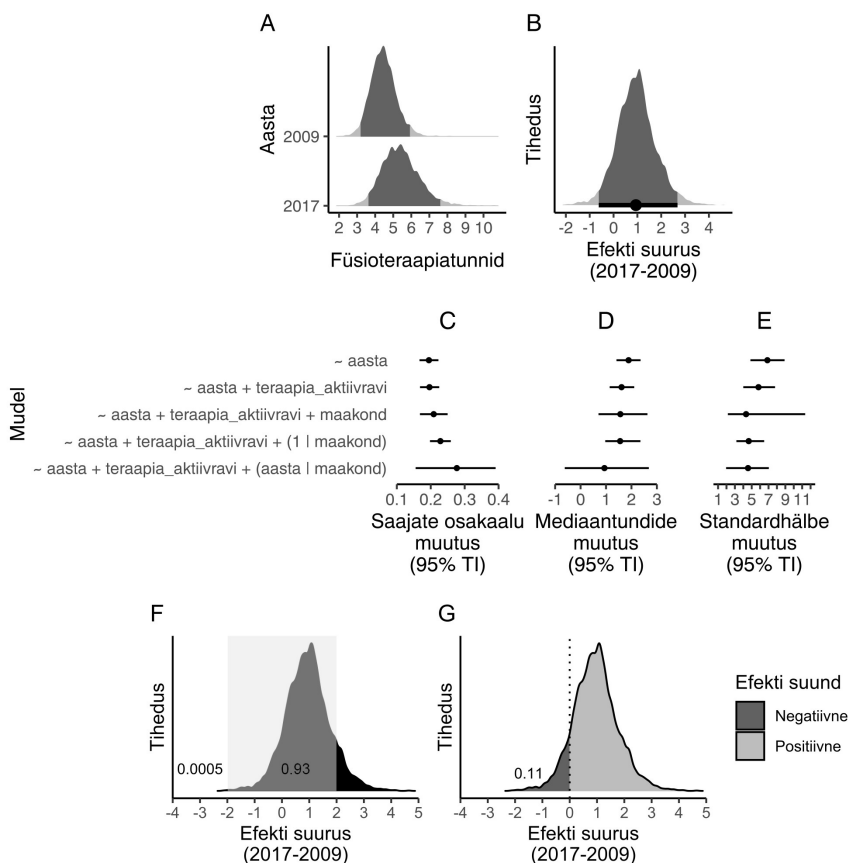
Joonis 1: Näide MCMC ahelate poolt mudeli $y \sim a + bx$ järeljaotusest võetud juhuvalim. (A) Numbriliselt on näidatud 3 andmepunkti 4000st. (B) Järeljaotuse valim 3D parameetriruumis. (C) 100-sammuline segment MCMC ahelatest igale parameetrile (tegelikult salvestati igale ahelale 1000 sammu). Siin on iseseisvalt jooksatud 4 MCMC ahelat, mille konvergens on hea. D. Parameetrite järeljaotused tihedusplotina.



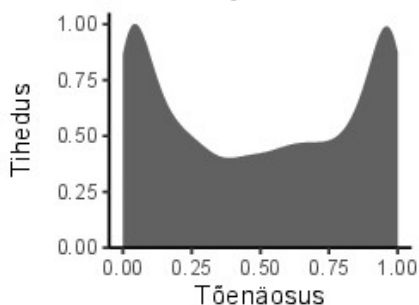
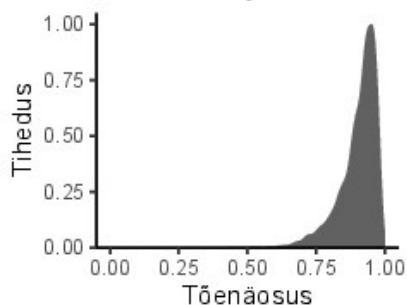
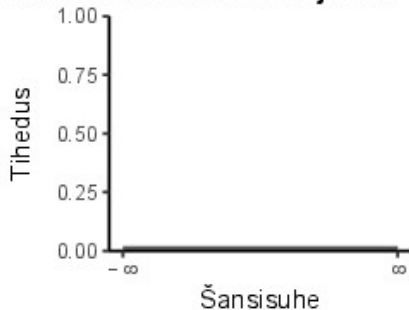
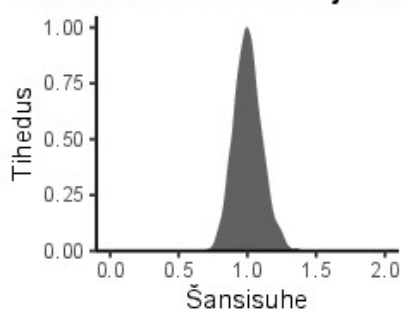
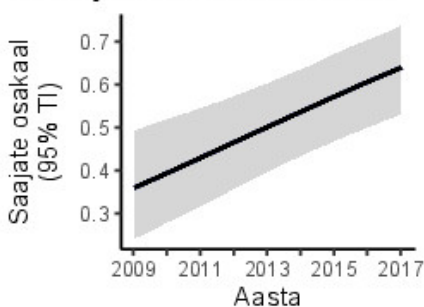
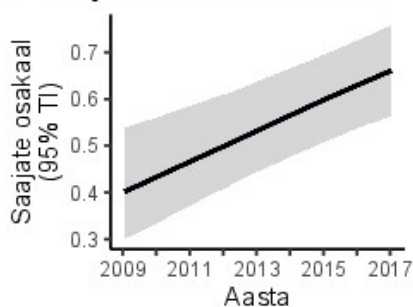
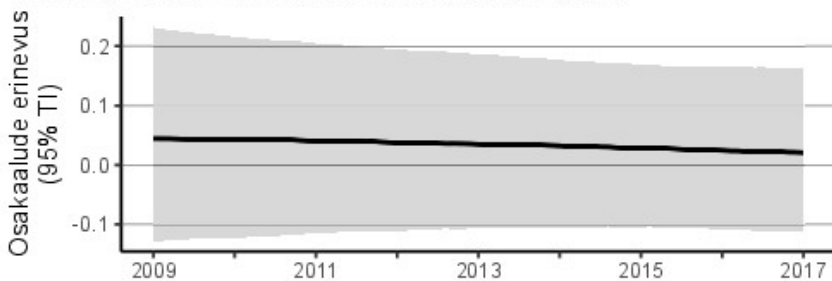
Joonis 2: Ajalised trendid kasutatud füsioteraapias (A) ja peamise tulemuslikkuse mõõdiku jaotus (B). Järe ravi mittesaanute osakaal küll väheneb, kuid mõistlikus hulgas ravi saanute (>10 tunni) osakaal jääb ikkagi alla viiendiku patsientidest, kes seda ravi vajavad. Peamine tulemuslikkuse mõõdiku – järe ravi ajal kasutatud füsioteraapia tundide – tunnuse jaotus on nulle täis ja tugevalt paremale kaldu. Positiivsete väärtuste standardhälbe (SD) vahemik on joonisel näidatud kui keskmine $(8,7) + 1SD$.



Joonis 3: Erinevate mudelite sobivuse hindamine y -tunnuse positiivsete väärtuste analüüsiks. (A) Mudel normaaljaotusel põhineva tõepärafunktsiooniga. (B) Mudel kaldus normaaljaotusel põhineva tõepärafunktsiooniga. (C) Mudel lognormaaljaotusel põhineva tõepärafunktsiooniga. Y – saadud ravitundide arvu empiiriline jaotus, Y_{rep} – mudelist genereeritud 50 valimi, millest igaüks on sama suur kui empiiriline valim, jaotused.



Joonis 4: Näide bayesaanliku regressioonimodeli järeljaotuse kasutamisest teaduslike tulemuste presenteerimiseks модели $tunnid \sim aasta + teraapia_{aktiivravi} + (aasta|maakond)$ põhjal. (A) Modelist ennustatud jaotused uuringu esimesele ja viimasele aastale. (B) Efekti suuruse järeljaotus, ehk viimase ja esimese uuringuaasta y -tunnuste vahe. (C) Erinevate sõltumatute tunnustega Bernoulli modelite efekti suurused. (D) Erinevate sõltumatute tunnustega lognormaalsete modelite efekti suurused teraapiatundide mediaanile. (E) Erinevate sõltumatute tunnustega lognormaalsete modelite efekti suurused füsioteraapiatundide varieeruvusele. (F) Järeljaotusest arvatud tõenäosus, et keskmiste ravitundide muutus 9 aasta jooksul on väiksem kui 2 tundi, ning et trend on negatiivne rohkem kui -2 tundi. (G) Bayesiaanlik p -väärtus, ehk tõenäosus, et ravitundide muutus trend on negatiivne.

A Vabaliikme eeljaotus**B Vabaliikme eeljaotus****C Tõusukoefitsendi eeljaotus****D Tõusukoefitsendi eeljaotus****E Saajate osakaalu muutus****F Saajate osakaalu muutus****G Kahe mudeli osakaalu muutuste erinevus**

Kirjandus

1. Ü. Maiväli. *Interpreting Biomedical Science*. Academic Press, 2015.
2. R. McElreath. *Statistical Rethinking; A Bayesian Course with Examples in R and Stan; Second Edition*. CRC Press, 2020.
3. A. Gelman, J. Hill, A. Vehtari. *Regression and Other Stories*. Cambridge University Press, 2020.
4. A. Hald. *A History of Parametric Statistical Inference from Bernoulli to Fisher*. Springer Science & Business Media, 2008.
5. P. Prommik, H. Kolk, Ü. Maiväli, M. Pääsuke, A. Märtsen. High variability in hip fracture post-acute care and dementia patients having worse chances of receiving rehabilitation: an analysis of population-based data from Estonia. *European Geriatric Medicine* 2020 Aug;11(4), 581–601.
6. P. Prommik, Ü. Maiväli, H. Kolk, A. Märtsen. Causal variation modelling identifies large inter- and intra-regional disparities in physical therapy offered to hip fracture patients in Estonia. *Disability and Rehabilitation* 2022 Aug;44(17), 4729–4737.
7. SD Team. *Stan Modeling Language Users Guide and Reference Manual, Version 2.27*. Kättesaadav: <https://mc-stan.org>
8. D. Bates, M. Mächler, B. Bolker, S. Walker. Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Softw.* **67** (2015), 1–48.
9. P.-C. Bürkner. Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal*. **10** (2018), 395–411.
10. B. Goodrich, J. Gabry, I. Ali, S. Brilleman. Bayesian applied regression modeling via Stan. Kättesaadav: <https://mc-stan.org/rstanarm>
11. S.L. Brilleman, E.M. Elçi, J.B. Novik, R. Wolfe. Bayesian Survival Analysis Using the rstanarm R Package. arXiv:2002.09633 [stat.CO]

12. J. Gabry, D. Simpson, A. Vehtari, M. Betancourt, A. Gelman. Visualization in Bayesian workflow. *J. Royal Stat. Soc. (series A)* **182** (2019), 389–402.
13. P. Bürkner. Parameterization of response distributions in brms. Kättesaadav: https://ftp.sun.ac.za/ftp/CRAN/web/packages/brms/vignettes/brms_families.html
14. M.J. Swat, P. Grenon, S. Wimalaratne. ProbOnto: ontology and knowledge base of probability distributions. *Bioinformatics* **32** (2016), 2719–2721.
15. SD Team. *Brief Guide to Stan's Warnings*. Kättesaadav: <https://mc-stan.org/misc/warnings.html>
16. A. Gelman, A. Vehtari, P.-C. Bürkner. Bayesian workflow. Mitteilmunud, kättesaadav: <http://www.stat.columbia.edu/gelman/research/>
17. N.P. Lemoine. Moving beyond noninformative priors: why and how to choose weakly informative priors in Bayesian analyses. *Oikos* **128** (2019), 912–928.
18. J. Pearl, M. Glymour, N.P. Jewell. *Causal inference in statistics*. Wiley, 2016.
19. N.N.N. Dongen van, J.B. Doorn van, Q.F. Gronau, D. Ravenzwaaij van, R. Hoekstra, M.N. Haucke, et al. Multiple Perspectives on Inference for Two Simple Statistical Scenarios. *The American Statistician*. **73**, 2019, 328–339.
20. Ü. Maiväli. Mis asi on tõenäosus ja miks me sellest hoolime. *Schola Biotheretica* XLVI, 2020.
21. J. Lember. Bayesi talumatu kergus. *Schola Biotheretica* XLVI, 2020.
22. A. Gelman, Y. Yao. Holes in Bayesian Statistics. Mitteilmunud, kättesaadav: <http://www.stat.columbia.edu/gelman/research/>

23. R. Schoot van de, S. Depaoli. Bayesian analyses: where to start and what to report. *The European Health Psychologist* **16** (2014), 75–84.
24. J. Doorn van, D. Bergh van den, U. Böhm, F. Dablander, K. Derks, T. Draws, et al. The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychonomic Bulletin & Review*. **28** (2021), 813–826.
25. T. Päll, H. Luidalepp, T. Tenson, Ü. Maiväli. A field-wide assessment of differential expression profiling by high-throughput sequencing reveals widespread bias. *PloS Biol.* 2023; Mar2;21(3): e3002007.
26. J. Pearl, D. Mackenzie. *The Book of Why*. Basic Books, 2018.