

# Bayesi talutav raskus: tõenäosuslik andmeanalüüs praktiku pilgu läbi

Pärt Prommik<sup>1,2,3</sup>, Aare Märtson<sup>1,2</sup>, Helgi Kolk<sup>1,2</sup>, Ülo Maiväli<sup>4</sup>

<sup>1</sup>Traumatoloogia ja ortopeedia kliinik, Kliinilise meditsiini instituut, Tartu Ülikool;

<sup>2</sup>Traumatoloogia ja ortopeedia kliinik, Tartu Ülikooli Kliinikum;

<sup>3</sup>Sportideaduste ja füsioteraapia instituut, Tartu Ülikool;

<sup>4</sup>Tehnoloogainstituut, Tartu Ülikool.

## Olulised mõisted

**Tõenäosus** – numbriline mõõt meie ebakindluse määrale parameetriväärtuse 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) – simulatsioonipõhine meetod, millega võetakse juhuvalim järeljaotusest ja moodustatakse Markovi ahel.

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

## Sissejuhatus

Andmeanalüüsi taga, kui see toimub teaduses, on enamasti soov lasta andmetel heita valgust teaduslikele teooriatele. See protsess on kahesuunaline, 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 hoomata 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, 23), milledest 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 punkthinnanguna 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 (Hald, 2008 Hald, A. (2008). *A History of Parametric Statistical Inference from Bernoulli to Fisher, 1713-1935*. Springer Science & Business Media.).

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 ennustada mudeli jääkidelt). Seda arvutust kutsutakse mudeli määramiseks (fittimiseks).

Bayesiaanlikus kuues näeb see mudel välja nii:

$$y \sim N(\mu, \sigma^2) \quad (1)$$

$$\mu = a + bx \quad (2)$$

Esimene rida annab tõepäramudeli, ehk ennustatava  $y$ -muutuja jaotusmudeli, mis ütleb, et  $y$ -muutujale omistatavad väärtused on määratud stohhastiliselt (tilde näitab stohhastilist seost) normaaljaotuse poolt, mille parameetrid on  $\mu$  ja  $\sigma$ , millest esimene on identsussuhtes  $y$ -i aritmeetilise keskmisega ja teine vastab  $y$ -i standardhälbele. Teine rida, nn protsessimudel, defineerib  $\mu$  ümber  $y$ -muutuja ja  $x$ -muutuja deterministliku sõltuvussuhtena. Seega kaob mudelist  $\mu$  ja selle asemele tulevad  $a$  ja  $b$  parameetrid. Seega ütleb täismudel, et  $y$ -muutuja jaotus (ja seega ka tõepäramudel) on tingimuslik  $x$ -muutuja väärtusest.

Meie mudelis määratakse kolm parameetrit ( $a$ ,  $b$  ja  $\sigma$ ) ja nõnda fikseeritud mudel on generatiivne, st sellest saab genereerida uusi  $y$ -muutuja väärtusi, mis sõltuvad vastavalt protsessimudelile  $x$ -muutujale omistatud väärtustest. Selline generatiivsus iseloomustab kõiki bayesiaanlikke mudeleid, kuid mitte kõiki sageduslikke analooge.

Bayesiaanlikus regressioonis vastab mudeli igale parameetrile eeljaotus - meie näites  $a$ -le,  $b$ -le ja  $\sigma$ -le, mille abil saame mudelisse lisada andmetest sõltumatut informatsiooni võimalike parameetriväärtuste kohta. Näiteks tõusukoefitsient  $b$  eeljaotuse tsentreerimine nullile lisab mudelile konservatiivsust, „tõmmates“  $b$  järeljaotust nulli suunas seda rohkem, mida kitsam on see eeljaotus ja mida vähem me mudelisse lisame valimiandmeid<sup>1</sup>. Mõistlikult valitud eeljaotused vähendavad ohtu, et me valimeväärtusi või kallutatud andmete põhjal teeme halbu teaduslikke järeldusi. Kui meil on piisavalt kvaliteetseid andmeid ja piisavalt lihtne mudel, 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 mudeli 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 tükudena, 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

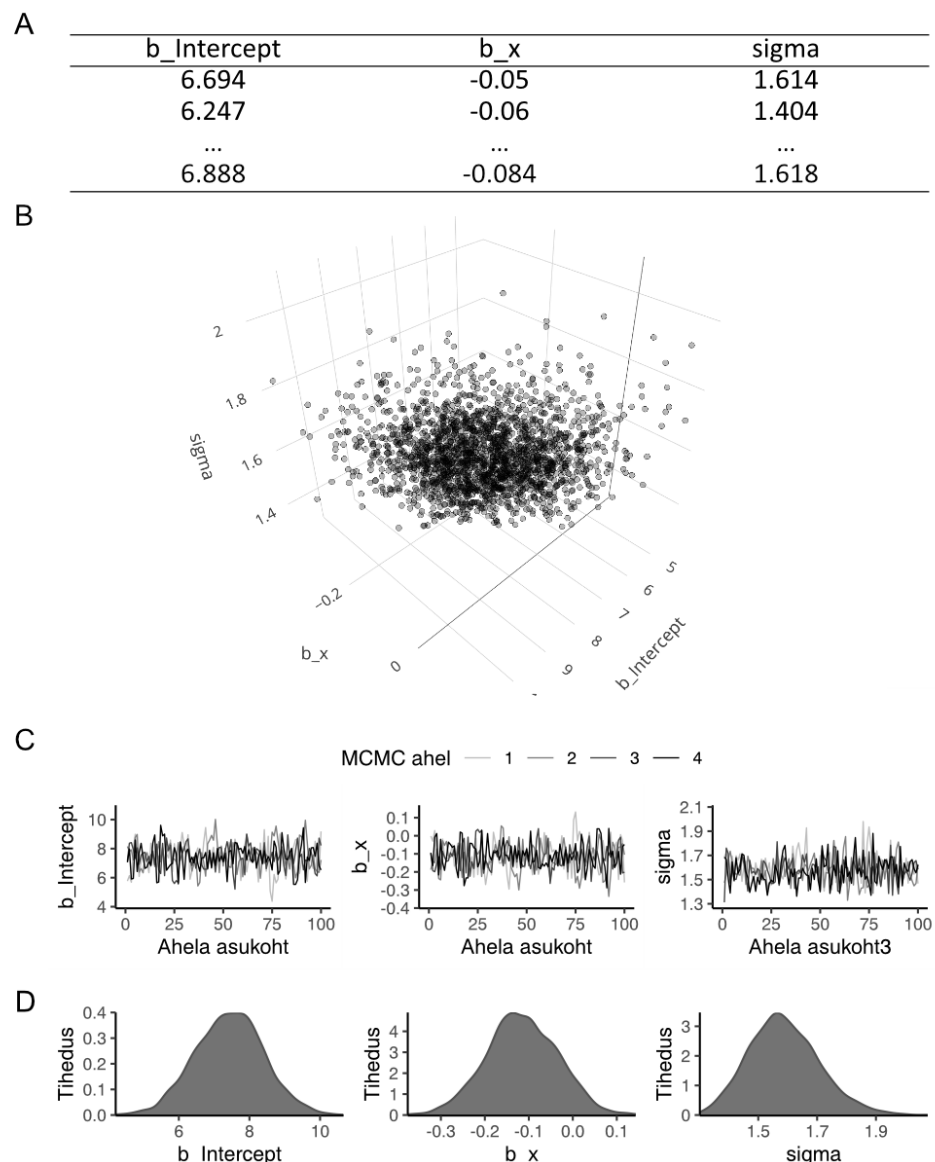
---

<sup>1</sup>Sageduslikus statistikas kutsutakse analoogset lähenemist *penalized likelihood*, bayesiaanlikus aga mudeli regulariseerimiseks.

trajektooridel ja „tõuk“ on ühtlaselt karvane, ilma siledate 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). Järeljaotustega (Joonis 1A), saab teha aritmeetilisi tehteid. Näiteks kui tahame teada, milline on  $y$ -muutuja ennustatav keskvärtus juhul kui  $x = 1$ , liidame  $a$  ja  $b$  koefitsientide järeljaotuste vektorid ( $a + b \cdot 1$ ), mis annab meile järeljaotuse sellele  $y$  keskvärtusele.

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 jooksnutud 4 MCMC ahelat, mille konvergens on hea. D. Parameetrite järeljaotused tihedusplotina.

Regressioonimudelit on lihtne laiendada (i) asendades tõepäramudelis normaaljaotuse mõne



muu tõenäosusjaotusega (binoomjaotusega, lognormaaljaotusega, jms), (ii) modifitseerides protsessi mudelit, k.a. mitte-linearseks, (iii) defineerides ümber rohkem kui ühe tõepäramudeli parameetri (näit  $\mu$  ja  $\sigma$ ), sama või erinevate protsessi mudelitega,

seome  $x$ -muutujaga varieeruvuse hinnangu, (iv) pannes mudelisse rohkem kui ühe ennustatava muutuja, (v) transformeerides ennustatava  $y$ -muutuja(d) näiteks logit-transformatsiooniga, (vi) lisades eeljaotustele meta-eeljaotusi muudame mudeli mitmetasemeliseks (2).

## Töövahendid

### Andmed ja uurimisküsimus

Bayeesanliku töövoog illustreerimiseks kasutame näidet, kus uurisime reieluu puusaliigese piirkonna murruga haigete ravikäsitlust Eestis (3,4). 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 järelravis. Ravi esimeses etapis (aktiivravis) haige hospitaliseeritakse, opereeritakse ja seejärel alustatakse füsioteraapiaga. Võrdlemisi lühikesele aktiivravile järgneb järelravi, 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

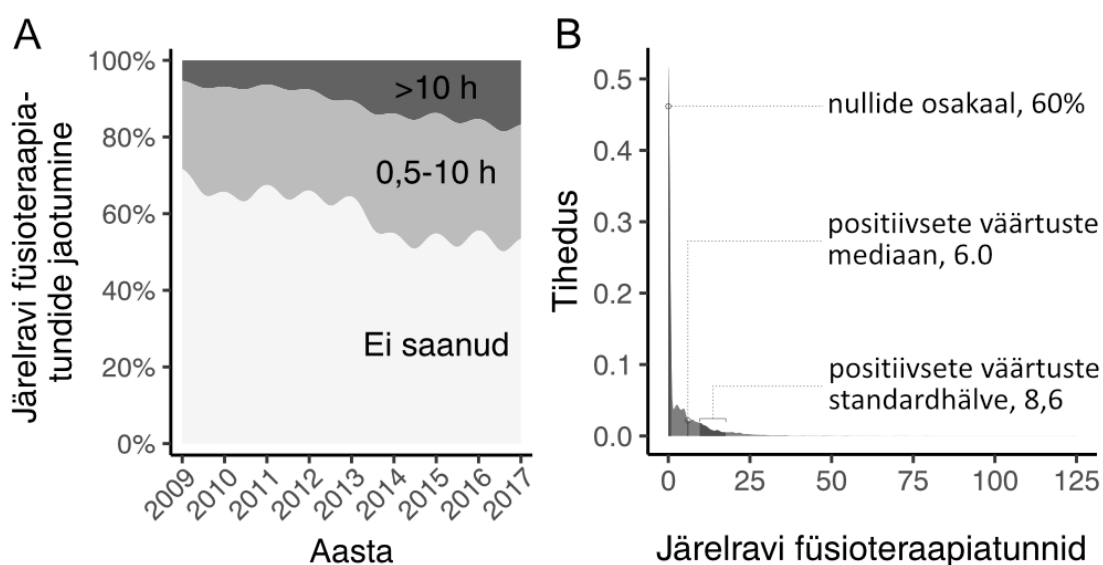
Bayesi teoreemi asemel määratakse mudeli koefitsiendid Markovi ahelate Monte Carlo (*Markov chain Monte Carlo*, MCMC) simulatsioonide abil, mis korjab juhuvalimi otse järeljaotusest. MCMC ahelate kasutamine on muudetud nii lihtsaks, et enamasti pole kasutajal vaja sellele palju mõelda. Siin omasid suurt rolli esimesed lihtsalt kasutatavad Markovi ahelaid jooksutavad programmid Bugs, winBugs ja JAGS, mille on praeguseks suuresti asendanud Stan, mis kasutab Hamiltonian Monte Carlo meetodit (5). Kuna Stan-i on keerulisevõitu õppida, kasutatakse mudelite ehitamiseks sageli abiprogramme, millest levinumad on R::rstanarm ja R::brms, mis võimaldavad mudeleid kodeerida R::lme4 ja sarnaste pakettide süntaksit kasutades (6–8). 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). Rstanarm keskendub levinumate sageduslike mudelite analoogidele (k.a. ANOVA ja elulemuse mudelid) (8,9). Brms annab laiemad valikud mudelite ehitamiseks ja on ka laiemas kasutuses.

# Bayesiaanlik töövoog

## 1. Esmane andmeanalüüs

Esmane ehk eksploratiivne 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 mudelite vahel (10).

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 (4), 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 regressioonianalüüsi vajalikkust.



Joonis 2. Ajalised trendid kasutatud füsioteraapias (A) ja peamise tulemuslikkuse mõõdiku jaotus (B). Järelravi mittesaanute osakaal küll väheneb, kuid mõistlikus hulgas ravi saanute (>10 tunni) osakaal jääb ikkagi alla viiendiku patsientidest, kes seda ravi vajavad. Peamine tulemuslikkuse mõõdiku – järelravi 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) + 1 SD.

## 2. Tõepärafunktsiooni leidmine

Bayesiaanlikus regressioonis ehitame oma mudelid - olgu need lineaarsed või mittelineaarsed, ühe- või mitmetasemelised - samale mõistelisele alusele ja jooksutame neid sama MCMC algoritmiga. Esmalt spetsifitseerime 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 (**Error! Reference source not found.2B**), siis analüüsime ajalisi trende kolmes osas: (i) füsioteraapiat saanute osakaal, (ii) saadud füsioteraapiatundide arv ja (iii) saadud füsioteraapiatundide varieeruvus patsientide vahel. Niisiis proovime leida kõigile analüüsi osadele sobivad tõepärafunktsioonid. Brms pakub mudeldamiseks 40 erinevat tõepäramudelit ning kasutajal on võimalik ka oma soovi järgi spetsifitseerida mudeli jaotuslik vorm. Kuna kõik bayesiaanlikud mudelid on generatiivsed (mudelist saab genereerida uusi andmepunkte), aitab sobivat tõepärafunktsiooni valida määratud mudelist *in silico* genereeritud andmete võrdlemine valimiandmetega (*posterior predictive check*). Kui võrdlus näitab halba kokkulangemist, siis võime muuta tõepärafunktsiooni, sõltumatuid tunnuseid, nendevahelisi seoseid või lisada mudelisse konditsionaalseid varieeruvusi. Sõltuvalt analüüsi eesmärgist, võib 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.

Esimene 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)
```

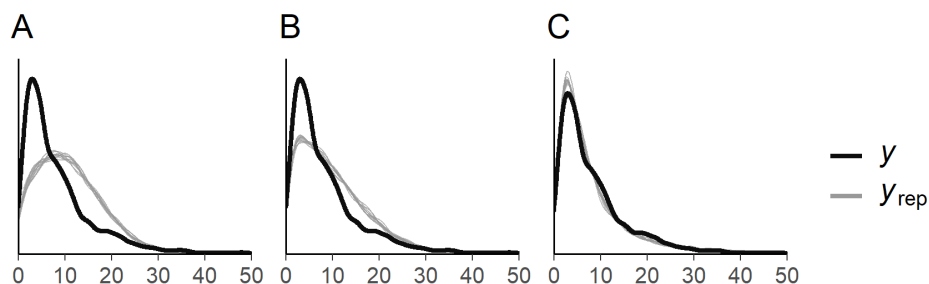
Analüüsi teine osa keskendus järelravi ajal saadud füsioteraapia mahu hindamisele. Mudelile sobiva jaotusliku vormi leidmiseks proovisime kolme tõepärafunktsiooni: normaaljaotus, kallutatud normaaljaotus ja lognormaaljaotus. Lisaks täiendame kõiki kolme mudeleid selliselt, et nad üheaegselt mudeldaksid 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 = „skew_normal“, data = andmed)
```

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

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



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,  $Y_{\text{rep}}$  – mudelist genereeritud 50 valimi, millest igaüks on sama suur kui empiiriline valim, jaotused.

Lõpetuseks, brms lubab meil mudelite ennustusi reparametriseerida (11), 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 (12). Käesolevas analüüsis reparametriseerime lognormaalse mudeli selliselt, et see annaks tulemused loomulikus skaalas ehk tundides: füsioteraapia tundide mahu mediaantundidena ja varieeruvuse standardhälbena<sup>2</sup>.

### 3. 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 ajalised 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:

$$\text{teraapia\_jarelravi} \sim 1$$

$$\text{teraapia\_jarelravi} \sim \text{aasta}$$

$$\text{teraapia\_jarelravi} \sim \text{aasta} + \text{teraapia\_aktiivravi}$$

$$\text{teraapia\_jarelravi} \sim \text{aasta} + \text{teraapia\_aktiivravi} + \text{maakond}$$

$$\text{teraapia\_jarelravi} \sim \text{aasta} + \text{teraapia\_aktiivravi} + (1 \mid \text{maakond})$$

$$\text{teraapia\_jarelravi} \sim \text{aasta} + \text{teraapia\_aktiivravi} + (\text{aasta} \mid \text{maakond})$$

<sup>2</sup>Reparametriseerimise kood on leitav Prommik et al. 2021 lisast.

Lognormaalsed mudelid mudeldasid paralleelselt ka standardhälvet, kasutades samu sõltumatuid tunnuseid. (footnote: mudeli kodeering brms-s: `brm(bf(teraapia_jarelravi ~ aasta + teraapia_aktiivravi + (aasta | maakond), sigma ~ aasta + teraapia_aktiivravi + (aasta | maakond)), family = „lognormal“)`) 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 posteriori 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 Gelman, et al. (2020) ja nende lahendamiseks on Stani arendusmeeskond välja andnud soovitusi (13). Mõnikord lahenevad mudeli konstrueerimisega seotud probleemid andmete transformeerimisega (näiteks y-tunnuse tsentreerimisel, standardiseerimisel, logaritmimeerimisel või logit-skaalasse transformeerimisel) (14). Teinekord tuleb muuta eeljaotus informatiivsemaks ehk rangemalt piirata mudeli erinevate parameetrite ulatusi või protsessimudel reparametriseerida (McElreath 2020).

#### 4. Mudelite võrdlemine

Parima mudeli valik sõltub analüüsi eesmärgist. Põhjusliku analüüsi puhul kasutame selleks suunatud atsüklilisi diagrame (*directed acyclic graph e 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. Tabelis 1 toome kuue mudeli bayeesaanliku *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 mudelid puhul on kaks 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 (sigma) 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 kinnituseks, et meie arusaam andmeid tootva protsessi hierarhilisest struktuurist on adekvaatne.

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

Mudel	Bernoulli		Lognormaalne	
	elpd_diff	se_diff	elpd_diff	se_diff



$\sim \text{aasta} + \text{teraapia\_aktiivravi} + (\text{aasta} \mid \text{maakond})^*$	0.0	0.0	0.0	0.0
$\sim \text{aasta} + \text{teraapia\_aktiivravi} + (1 \mid \text{maakond})^*$	0.0	0.0	-32.9	8.9
$\sim \text{aasta} + \text{teraapia\_aktiivravi} + \text{maakond}^*$	-0.1	0.5	-33.3	9.2
$\sim \text{aasta} + \text{teraapia\_aktiivravi}^*$	-492.6	30.9	-308.7	25.7
$\sim \text{aasta}^*$	-521.9	31.7	-342.7	27.3
$\sim 1^*$	-611.4	34.1	-376.1	28.3

\*samu sõltumaid tunnuseid kasutati standardhälbe paralleelseks mudeldamiseks lognormaalsetes 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 järeljaotusele;

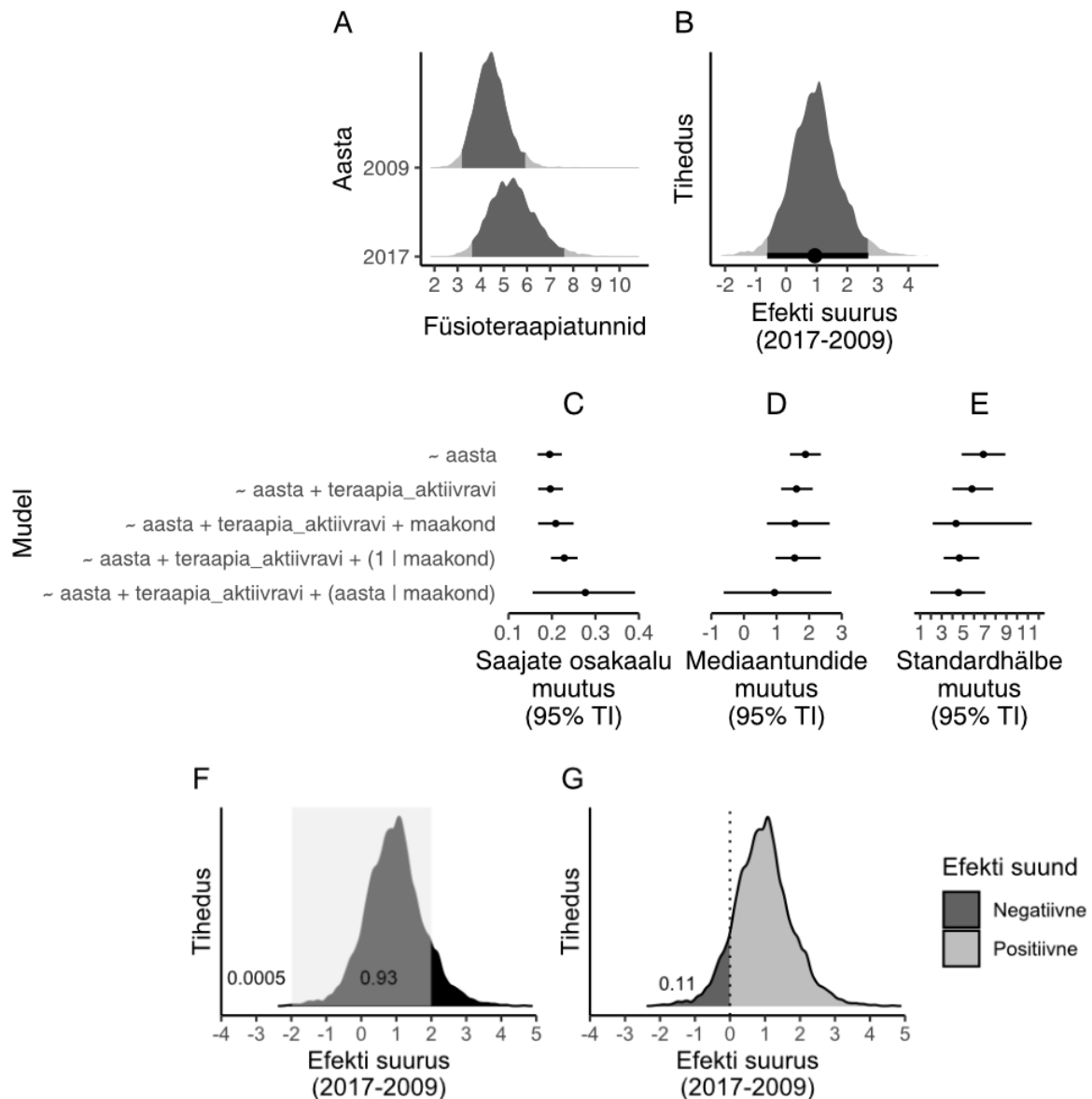
## 5. Järeljaotuste informatiivsus

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 statistiku, eesmärgiga kirjeldada taastusravi ajaliste trende.

Me saame järeljaotust kirjeldada kasutades tavapäraseid keskmist iseloomustavaid statistiku (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 ( $ES = \hat{Y}_{2017} - \hat{Y}_{2009}$ ) 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.



Joonis 4. Näide bayesaanliku regressioonimudeli järeljaotuse kasutamisest teaduslike tulemuste presenteerimiseks mudeli  $tunnid \sim aasta + teraapia\_aktiivravi + (aasta | maakond)$  põhjal. (A) Mudelist ennustatud jaotused uuringu esimesele ja viimasele aastale. (B) Efekti suuruse järeljaotus, ehk viimase ja esimese uuringuaasta y-tunnuste vahe. (C) Erinevate sõltumatute tunnustega Bernoulli mudelite efekti suurused. (D) Erinevate sõltumatute tunnustega lognormaalsete mudelite efekti suurused teraapiatundide mediaanile. (E) Erinevate sõltumatute tunnustega lognormaalsete mudelite 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 muutuse trend on negatiivne.

## 6. 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 1/6 ravitundide varieeruvusest seletatav maakonna mõjuga ravitundidele ja ülejäänud 5/6 jääb populatsiooni tasemele seletust ootama. Seevastu varieeruvuse tõus ajas jaguneb üsna võrdselt populatsiooni taseme ( $sigma\_aasta\_tsentreeritud$ ) ja maakonna taseme ( $sd(sigma\_aasta\_tsentreeritud)$ ) 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 ligikaudne 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,  $cor(Intercept, aasta\_tsentreeritud)$  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			
Intercept	1,937698	1,750588	2,126124
sigma_Intercept	1,948379	1,724296	2,175053
aasta_tsentreeritud	0,042114	0,013005	0,067134
teraapia_aktiivravi_tsentreeritud	0,042729	0,028986	0,058896
sigma_aasta_tsentreeritud	0,079122	0,035086	0,120775

sigma_teraapia_aktiivravi_tsentreeritud	0,042846	0,020141	0,070378
Maakonna tasemel koefitsiendid			
sd(Intercept)	0,338894	0,223185	0,531837
sd(sigma_Intercept)	0,407452	0,270199	0,626901
sd(aasta_tsentreeritud)	0,041453	0,021766	0,072774
sd(sigma_aasta_tsentreeritud)	0,066316	0,03286	0,115814
cor(Intercept,aasta_tsentreeritud)	0,178179	-0,41724	0,674109
cor(sigma_Intercept,sigma_aasta_tsentreeritud)	-0,35149	-0,81366	0,283575

## 7. Eeljaotuse mõju mudelile

Bayesaanlike mudelite määramine sõltub mitte ainult andmetest (tõepärafunktsiooni kaudu), vaid ka eeljaotusest. Range ehk kitsas eeljaotus võib mudeli parameetri järeljaotust 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 ruumi kitsendades selle algoritmilist määramist MCMC meetoditega - eriti väheste andmete korral (Lemoine 2019, Gelman, et al. 2020).

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 vaikesätteid, sest meil on suur ja väga kvaliteetne andmestik<sup>3</sup>. Range eeljaotuse mõju demonstreerimiseks võrdleme kahte versiooni sama protsessimudeliga hierarhisest Bernoulli mudelist

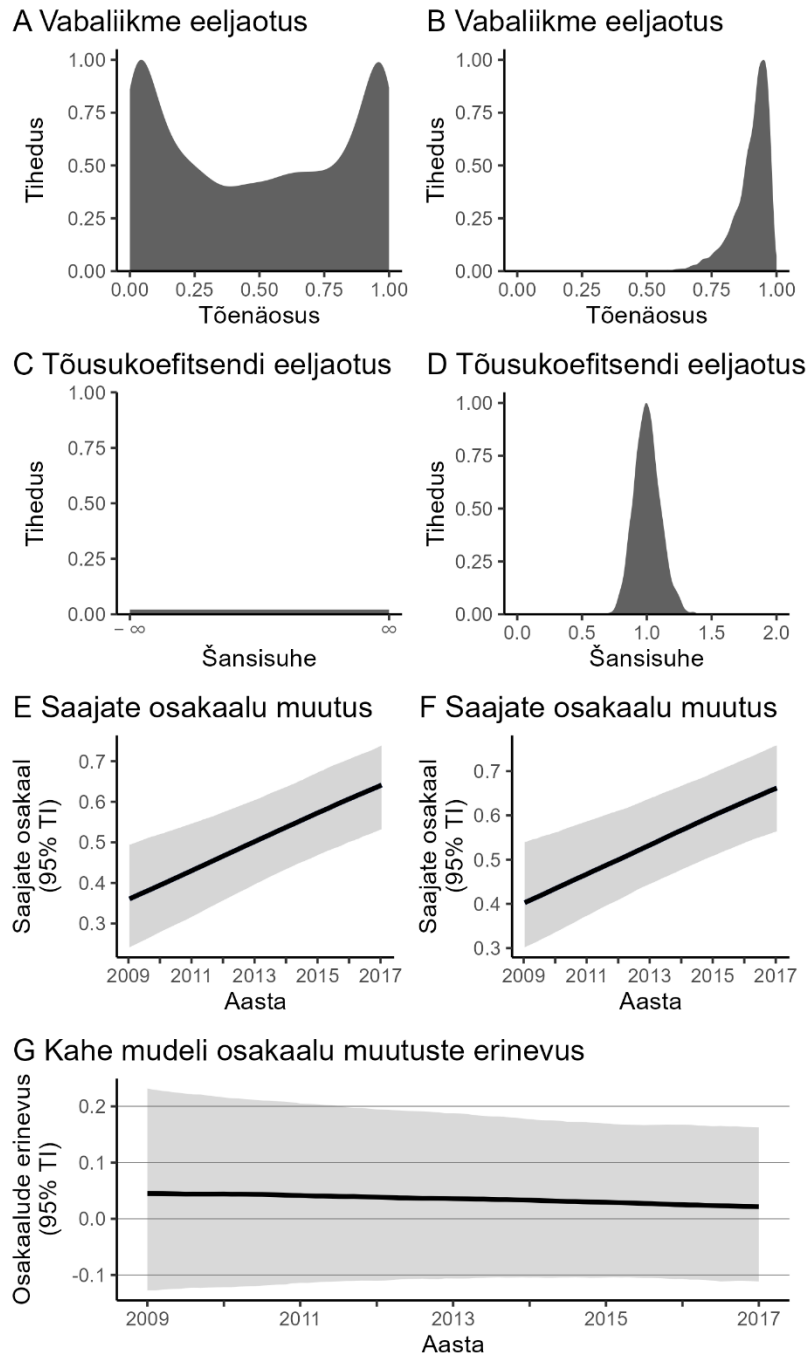
```
brm(teraapia_jarelravi_binaarne ~ aasta + teraapia_aktiivravi + (aasta | maakond), family = „bernoulli“, link = „logit“, data = andmed).
```

Ühe neist määrame brms-i vaike-eeljaotusega ja teisele anname range ja ebamõistliku eeljaotuse, mis ilustab 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

<sup>3</sup>Brms-i eeljaotuste vaikesätted ei mõjuta enamasti tõusukoefitsiente (nende eeljaotused on tasased) ja mõjutavad väga nõrgalt ülejäänud mudeli parameetreid.

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.



Joonis 5. Vaike-eeljaotusega (A, C, E) ja informatiivsete eeljaotusega (B, D, F) mudelite võrdlus. A-D näitavad kasutatud eeljaotuseid. E-F näitava mudelite ennustusi vaadeldud 9-aastasele perioodile. G võrdleb kahe mudeli ennustuste erinevusi vaadeldud perioodil (informatiivse eeljaotusega mudeli ennustus miinus vaike-eeljaotusega mudeli ennustus).

## 8. Teaduslike järelduste tegemine

Rakendusliku statistika üks peamisi eesmärke 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 (16). 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 tõepärafunktsioonid ja mudelite determinismliku osa struktuurid, mis võimaldavad realistlikke mudelipõhiseid ennustusi. Mitme mudeli tulemuste üheskoos vaatlemine andis parema arusaama tegeliku mõju suunast ja suurusest. 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 maakondade vahel kõigis kolmes näitajas, ehk seda, et nii lisandunud raviressurs kui selle kasutamise efektiivsus liikus kohalikele tasemele väga ebaühtlaselt (vt [Pronnik jt 2021](#) 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 (17): (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 järeлгаotused 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 nullhüpoteesi kehtimise korral (18).

Bayesiaanlikul lähenemisel on leitud olema ka puuduseid, sealhulgas sügavalt teoreetilisi (19,20). Meie arvates on praktilist laadi puudused, mis reaalselt mõjutavad teadlast, kes rakendab bayesiaanlikku töövoogu, järgmised:

- (i) Puudub sisse töötatud vorm bayesiaanlike tulemuste avaldamiseks. Kuigi leidub publikatsioone, mis pakuvad reegleid selleks, lonkab siiani nende ühtne rakendamine praktikas (21,22). Siin aitab analüüsikoodi täismahus avaldamine artikli lisana, soovitatavalt koos lühikese õpetusega, kuidas seda koodi kasutada ja selles olevaid mudeliobjekte avada (vt näiteks Päll et al. 2021).
- (ii) Informatiivsete eeljaotuste defineerimine keerukamatele mudelitele võib osutuda keerukaks (vt joonist 6A). Selle probleemiga tegeletakse aktiivselt.
- (iii) 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 (Pearl 2018).
- (iv) Ajakirjade retsensendid kipuvad võtma bayesiaanlikku andmeanalüüsi kriitikavabalt, mis langetab kokkuvõttes avaldatud tööde kvaliteeti, või mis veel hullem, tegema bayesiaanlikele töödele sageduslikku kriitikat, mis raskendab avaldamist, lisamata midagi kvaliteedile.
- (v) Bayesiaanlikud mudelid jooksevad aeglaselt. Meie keerulisemate mudelite läbiarvutamine võtab sülearvutil kuni paar tundi. Lihtsad mudelid jooksevad tavaliselt sekundite kuni minutite jooksul.

## Viited

- 1 Maiväli Ü. Interpreting Biomedical Science - 1st Edition. Cambridge, Massachusetts, United States: Academic Press, 2015.
- 2 McElreath R. Statistical Rethinking: A Bayesian Course with Examples in R and STAN. CRC Press, 2020.



- 3 Prommik P, Kolk H, Maiväli Ü, Pääsuke M, Märtsen A. 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. *Eur Geriatr Med* 2020;11:581–601.
- 4 Prommik P, Maiväli Ü, Kolk H, Märtsen A. Causal variation modelling identifies large inter- and intra-regional disparities in physical therapy offered to hip fracture patients in Estonia. *Disabil Rehabil* 2021;doi: <https://doi.org/10.1080/09638288.2021.1918772>.
- 5 Stan Development Team. Stan Modeling Language Users Guide and Reference Manual, Version 2.27. Stan Model Lang Users Guide Ref Man 2021. <https://mc-stan.org> (accessed July 11, 2021).
- 6 Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. *J Stat Softw* 2015;67:1–48.
- 7 Bürkner P-C. Advanced Bayesian multilevel modeling with the R package brms. *R J* 2018;10:395–411.
- 8 Goodrich B, Gabry J, Ali I, Brilleman S. Bayesian applied regression modeling via Stan. R package version 2.21.1 2020. <https://mc-stan.org/rstanarm>.
- 9 Brilleman SL, Elci EM, Novik JB, Wolfe R. Bayesian survival analysis using the rstanarm R package. *ArXiv200209633 Stat* 2020.
- 10 Gabry J, Simpson D, Vehtari A, Betancourt M, Gelman A. Visualization in Bayesian workflow. *J R Stat Soc Ser A Stat Soc* 2019;182:389–402.
- 11 Paul Bürkner. Parameterization of response distributions in brms 2020. [https://cran.r-project.org/web/packages/brms/vignettes/brms\\_families.html](https://cran.r-project.org/web/packages/brms/vignettes/brms_families.html) (accessed May 10, 2021).
- 12 Swat MJ, Grenon P, Wimalaratne S. ProbOnto: ontology and knowledge base of probability distributions. *Bioinformatics* 2016;32:2719–21.
- 13 Stan Development Team. Brief Guide to Stan's Warnings 2020. <https://mc-stan.org/misc/warnings.html> (accessed July 13, 2021).
- 14 Gelman A, Vehtari A, Simpson D, et al. Bayesian workflow. *ArXiv201101808 Stat* 2020.
- 15 Lemoine NP. Moving beyond noninformative priors: why and how to choose weakly informative priors in Bayesian analyses. *Oikos* 2019;128:912–28.
- 16 Pearl J, Glymour M, Jewell N. Causal Inference in Statistics: A Primer. 2016.
- 17 Dongen NNN van, Doorn JB van, Gronau QF, et al. Multiple Perspectives on Inference for Two Simple Statistical Scenarios. *Am Stat* 2019;73:328–39.
- 18 Maiväli Ü. Mis asi on tõenäosus ja miks me sellest hoolime. *Sch. Biotheretica XLVI*, 2020.
- 19 Lember J. Bayesi talumatu kergus. *Sch. Biotheretica XLVI*, 2020.
- 20 Gelman A, Yao Y. Holes in Bayesian Statistics. *J Phys G Nucl Part Phys* 2021;48:014002.
- 21 Schoot R, Depaoli S. Bayesian analyses : where to start and what to report. *Undefined* 2014.
- 22 van Doorn J, van den Bergh D, Böhm U, et al. The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychon Bull Rev* 2021;28:813–26.
- 23 Gelman A, Hill J, Vehtari A. Regression and Other Stories. Cambridge University Press, 2020.
- 24 Päll T, Luidalepp H, Tenson T, Maiväli Ü. A field-wide assessment of differential high throughput sequencing reveals widespread bias. 2021.