

Valkosolupitoisuuksien bayesilainen mallintaminen lasten leukemian ylläpito-hoidossa

Tilastotieteen pro gradu -tutkielma

06.06.2018
Santeri Karppinen
Matematiikan ja tilastotieteen laitos
Jyväskylän yliopisto

JYVÄSKYLÄN YLIOPISTO

Matematiikan ja tilastotieteen laitos

Santeri Karppinen: Valkosolupitoisuuksien bayesilainen mallintaminen lasten leukemian ylläpito-hoidossa

Tilastotieteen pro gradu -tutkielma, 56 sivua, ei liitteitä
06.06.2018

Tiivistelmä

Lasten akuutin lymfoblastileukemian ylläpitovaiheen hoidossa tehtävät lääkeannostuspäätökset pohjataan nykyisin potilaan veren valkosolupitoisuuteen, joka on hoidon tehokkuudesta kertova tekijä. Potilaalle sopiva lääkeannostus on hoidon onnistumisen ja turvallisuuden kannalta tärkeä, mutta sen löytäminen on vaikeaa, sillä annettu lääkitys näkyy valkosolupitoisuudessa viiveellä, ja potilaiden elimistön reagointi lääkitykseen on yksilöllistä. Sopivan lääkeannostuksen löytämistä hankaloittavat myös hoidonaikaiset tulehdukset, jotka voivat muuttaa valkosolupitoisuutta hetkellisesti.

Työ käsittelee akuuttiin lymfoblastileukemiaan sairastuneiden suomalaisten potilaiden veren valkosolupitoisuuden mallintamista kahdella bayesilaisella mallilla, differentiaaliyhtälömallilla sekä epälineaarilla tila-avaruusmallilla, jotka pohjautuvat kirjallisuudessa esitettyihin malleihin. Työssä esitellään ja laajennetaan kirjallisuudessa ehdotettuja malleja. Mallien estimointi tehdään suomalaisista potilaista koostuvaa aikasarja-aineistoa hyödyntäen. Tila-avaruusmalleja sekä potilaan tulehdusarvoa hyödynnetään ensimmäistä kertaa ilmiön mallintamisessa. Potilaan hoidonaikainen tulehdusarvo havaitaan hyödylliseksi muuttujaksi valkosolupitoisuuteen liittyvän epävarmuuden mallintamisessa.

Avainsanat: akuutti lymfoblastileukemia, valkosolupitoisuus, tulehdusarvo, bayesilainen differentiaaliyhtälömalli, bayesilainen epälineaarinen tila-avaruusmalli, laajennettu Kalman-suodin, adaptiivinen MCMC-algoritmi

Sisältö

1	Johdanto	1
2	Lasten ALL	2
2.1	Yleistä	2
2.2	Ylläpitovaiheen hoito	2
2.3	6MP:n yksilöllinen vaikutus leukosyyttipitoisuuteen	4
3	Aineisto	5
4	Kirjallisuudessa esitetyt mallit	8
4.1	Friberg ym. [2002] malli syöpälääkkeiden aiheuttamalle myelosupressiolle	8
4.2	Jayachandran ym. [2014, 2015] leukopoeesimalli ja 6MP:n farmakokineettinen malli	10
4.2.1	Leukopoeesimalli	10
4.2.2	6MP:n farmakokineettinen malli	13
5	Bayesilainen differentiaaliyhtälömalli	16
5.1	Mallin kuvaus	16
5.2	Parametrien priorijakaumat ja kiinnitetyt arvot	18
5.3	Estimointi	19
5.4	Konvergenssi	20
5.5	Potilaiden posteriorijakaumat	21
5.6	Mallin mukaiset leukosyyttipitoisuudet	23
6	Bayesilainen epälineaarinen tila-avaruusmalli	27
6.1	Mallin kuvaus	27
6.1.1	Yksinkertaistettu 6MP:n farmakokineettinen malli	28
6.1.2	Malli tulehdusarvolle	29
6.1.3	Tulehduksen liittäminen leukopoeesimalliin	30
6.1.4	Havaintoyhtälö	31
6.2	Parametrien priorijakaumat ja kiinnitetyt arvot	31
6.3	Estimointi	35
6.4	Konvergenssi	37
6.5	Potilaiden posteriorijakaumat	37
6.6	Mallin mukaiset leukosyyttipitoisuudet	40
7	Yhteenveto ja pohdintaa	44
7.1	Bayesilainen differentiaaliyhtälömalli ja Jayachandran ym. [2014, 2015] mallinnus	44
7.2	Pohdintaa bayesilaisesta epälineaarista tila-avaruusmallista	46
7.2.1	Mallin kontribuutio ja vertailu bayesilaiseen differentiaaliyhtälömalliin	46
7.2.2	Mallin kehitysideoita	47
7.2.3	Mallin hyödyntäminen	51

Työssä käytettävät lyhenteet

6MP 6-merkaptopuriini

ALL Akuutti lymfoblastileukemia

AZA Atsatiopriini

MeMP Metyylimerkaptopuriini

MTX Metotreksaatti

TGN 6-tioguaaniini

TGN_{DNA} 6-tioguaaniinin DNA:han sitoutunut deoksimuoto

TGN_{RBC} Punasoluihin sitoutunut 6-tioguaaniini

TPMT Tiopuriinimetyylitransferaasi

1 Johdanto

Akuutti lymfoblastileukemia (ALL) on yleisin lasten syöpä, johon sairastuu Suomessa vuosittain noin 40 lasta [Lohi ym., 2013]. Tässä työssä keskitytään ALL:n hoidon viimeiseen vaiheeseen, jota kutsutaan ylläpitovaiheeksi. Ylläpitovaiheessa tehtävät onnistuneet päätökset lääkeannostuksen suuruudesta vaikuttavat ALL:stä paranemiseen ja ovat yhteydessä syövän uusiutumattomuuteen. Nykyisin lääkeannostuspäätösten perustana on potilaan hoidonaikainen veren valkosolupitoisuus, jolla on yhteys hoidon tehokkuuteen. Lääkeannostuspäätösten tekeminen on vaikeaa, sillä ylläpitovaiheen lääkehoitojen vaikutus näkyy potilaan veren valkosolupitoisuudessa viiveellä ja riippuu potilaan elimistön yksilöllisestä reagoinnista annettaviin lääkkeisiin. Tärkein ylläpitovaiheen aikana käytettävä lääke on 6-merkaptopuriini. [Schmiegelow ym., 2014]

Tässä työssä tutkitaan kahta bayesilaista mallia, joilla ALL-potilaiden veren valkosolupitoisuuksia voidaan mallintaa ylläpitovaiheen aikana. Mallit sovitetaan suomalaisia ALL-potilaita sisältävään aineistoon. Mallien tutkimisen taustalla on pyrkimys kehittää ennustemalli, jolla potilaiden valkosolupitoisuuden todennäköisyysjakauma voidaan ennustaa ottaen huomioon potilaan elimistön yksilöllinen reagointi 6-merkaptopuriinilääkitykseen. Ennustemallin avulla potilaan lääkeannostuspäätösten tekeminen helpottuisi, sillä lääkityksen suuruus voitaisiin määrittää hoidossa tavoitteena olevaa leukosyyttipitoisuutta ennustavaksi. Tällöin hoidon tavoitteet saavutettaisiin tehokkaammin kuin seuraamalla suuntaa-antavia, kaikille potilaille yhteisiä hoito-ohjeita, joihin lääkeannostuspäätökset nykyään perustuvat. Ennustemallin hyödyntäminen päätöksenteossa voisi vähentää myös eri lääkäreiden subjektiivisesta näkemyksestä johtuvaa hoidon vaihtelua.

Kappaleessa 2 kuvaillaan lasten ALL:ää, ylläpitovaiheen hoitoa sekä 6-merkaptopuriinin yksilölliseen vaikutukseen liittyviä tekijöitä. Kappale 3 esittelee ALL-potilaiden ylläpitovaiheen hoidon ajalta kerätyn aikasarja-aineiston, johon työn mallit sovitetaan. Kappaleessa 4 esitellään Friberg ym. [2002] ja Jayachandran ym. [2014, 2015] elimistön valkosolutuotantoa sekä 6-merkaptopuriinin farmakokinetiikkaa¹ kuvaavat mallit. Nämä lähteet muodostavat pohjan tässä työssä tehtävälle mallinnukselle.

Kappaleet 5 ja 6 muodostavat työn ytimen, joka koostuu mallien ja mallinnustulosten esittelystä. Lähestymistapa mallintamiseen on työssä bayesilainen ja mallinnuksen tuloksia tarkastellaan tutkimalla potilaskohtaisia posteriorijakaumia sekä mallien sovitteita potilasaineistoihin. Kappaleessa 5 tarkastelun kohteena on differentiaaliyhtälömalli, joka on lähes suora sovellus Jayachandran ym. [2014, 2015] kehittämästä mallista. Toinen työn malleista on kappaleessa 6 esiteltävä epälineaarinen tila-avaruusmalli, joka kehittää Friberg ym. [2002] mallia huomioimalla potilaan hoidonaikaiset tulehdukset. Lisäksi mallissa 6-merkaptopuriinin farmakokinetiikan mallintamista on yksinkertaistettu Jayachandran ym. [2015] malliin verrattuna. Kappaleessa 6 esitettävät kirjallisuuden mallien parannukset ovat tämän työn keskeisin kontribuutio ALL:n ylläpitovaiheen valkosolupitoisuuksien mallintamiseen. Työ päättyy mallinnustulosten sekä kappaleen 6 mallin kehittämisen ja hyödyntämisen pohdintaan kappaleessa 7.

¹ Farmakokinetiikassa tutkitaan, mitä lääkeannokselle tapahtuu elävässä organismissa [Meibohm ja Dendorf, 1997].

2 Lasten ALL

2.1 Yleistä

Leukemiassa eli verisyövässä uusien verisolujen tuotannosta vastaavat luuytimen solut eivät toimi normaalisti. Epänormaalin verisolutuotannon seurauksena syntyvät solut² ovat viallisia ja kertyvät luuytimeen sekä vereen. Samalla normaalisti toimivien verisolujen tuotanto vähenee, mikä aiheuttaa ongelmia elimistössä sekä hoitamattomana kuoleman [VSSHP, 2017]. ALL:ssä potilaan verisolutuotanto tuottaa liikaa varhaisimusoluja eli lymfoblasteja tai B-lymfosyyttejä ja T-lymfosyyttejä, jotka ovat elimistön valkosoluihin eli *leukosyytteihin* kuuluvia soluja [National Cancer Institute, 2018]. ALL:ää esiintyy enimmäkseen lapsilla ja nuorilla, sillä 60 % potilaista on alle 20-vuotiaita. Lapsilla taudin esiintyvyys on suurinta 2 – 5 ikävuoden välillä. [Inaba ym., 2013]

ALL:n hoito on kehittynyt viime vuosikymmenten aikana paljon ja nykyään viiden vuoden tauditon elossaolo³ on lapsilla erittäin hyvä, lähes 90 % [Inaba ym., 2013]. Hoidon tarkoituksena on elimistön kaikkien syöpäsolujen tuhoaminen ja normaalin verentuotannon palauttaminen. Suomessa käytetyn NOPHO-ALL 2008 -hoito-ohjelman potilaat jaotellaan tavanomaisen, lisääntyneen ja korkean riskin ryhmiin [Lohi ym., 2013]. Jako riskiryhmiin tehdään taudin diagnosointiin liittyvien riskikriteerien sekä alkuvaiheen hoidon onnistumisen perusteella (NOPHO [2011], s. 34 – 39). Tavanomaisen ja lisääntyneen riskin ryhmissä hoito jakautuu neljään vaiheeseen, joita kutsutaan alkuhoito-, konsolidaatio-, tehostus- ja ylläpitovaiheiksi. Korkean riskin ryhmän hoito eroaa tavanomaisen ja lisääntyneen riskin ryhmien hoidosta siten, että alkuhoidon jälkeinen hoito perustuu ylläpitovaiheeseen asti noin kuukauden välein annettaviin solunsalpaajakuureihin (Lohi ym. [2013], NOPHO [2011], s. 56 – 96). Alkuhoitovaiheen tarkoituksena on saavuttaa remissio, joka tarkoittaa syöpäsolujen määrän merkittävää vähenemistä sekä normaalin verentuotannon palautumista [Lohi ym., 2013]. Konsolidaatiovaiheessa syöpäsolujen määrää pyritään edelleen vähentämään, sillä remissiosta huolimatta elimistössä on yleensä jäljellä vähäinen määrä syöpäsoluja. Tehostusvaihe on alkuhoitovaiheen kaltainen konsolidaatiovaiheen ja ylläpitovaiheen välille sijoittuva hoitojakso.

Tämän työn sisällön kannalta hoitovaiheista oleellisin on ylläpitovaihe, jonka aikana potilasta lääkitään suun kautta annettavilla solunsalpaajilla. Ylläpitovaihe on hoitovaiheista pisin, ja se kestää yleensä kunnes taudin diagnoosista on kulunut 2 – 3 vuotta [Schmiegelow ym., 2014]. ALL-potilaita hoidetaan ylläpitovaiheessa päivittäin annetulla 6-merkaptopuriinilla (6MP) sekä viikottain annetulla metotreksaatilla (MTX). MTX on solunsalpaaja, jonka käyttäminen 6MP:n kanssa lisää 6MP:n hyötyosuutta⁴ [Balis ym., 1987].

2.2 Ylläpitovaiheen hoito

Nykyisin ylläpitovaiheen hoito järjestetään yleensä siten, että hoidon aikana potilas käy lääkärissä kontrollikäynneillä tietyin väliajoin, tavallisesti yhden tai kahden viikon välein. Kontrollikäynneillä potilaalta otetaan verinäyte ja tehdään tarvittaessa muita tutkimuk-

² Tällaisia soluja kutsutaan leukemiassa syöpäsoluiksi.

³ Viiden vuoden tauditon elossaolo tarkoittaa osuutta potilaista, jotka eivät ole kuolleet eivätkä saaneet komplikaatioita viiden vuoden kuluessa hoidon päättymisestä.

⁴ Hyötyosuudella tarkoitetaan osuutta, joka lääkeaineesta pääsee verenkiertoon.

sia, joiden perusteella lääkäri tekee päätöksen siitä, minkälaisella lääkannoksella hoitoa jatketaan.

Eri hoito-ohjelmissa tällaisten lääkannostuspäätösten tukena käytetään eri kriteereitä. Yleisin päätöksenteossa seurattava muuttuja on potilaan veren leukosyyttipitoisuus, sillä solunsalpaajana 6MP hillitsee leukosyyttien tuotantoa ja siten hoidon tehokkuuden voidaan ajatella näkyvän potilaan hoidon aikaisissa leukosyyttipitoisuuksissa. Tutkimuksissa on havaittu, että potilaiden hoidon aikaiset matalat veren leukosyyttipitoisuudet ovat yhteydessä harvinaisempaan syövän uusiutumiseen sekä korkeampiin 6MP:n solumyrkyllisten metaboliittien tasoihin verisoluihin [Schmiegelow ym., 2014]. Solumyrkyllisten metaboliittien ilmaantuminen verisoluihin on tärkeää siksi, että metaboliitit tuhoavat syöpäsoluja.

Leukosyyttipitoisuuksille on määritetty tavoitepitoisuuksia, jotka vaihtelevat hieman hoito-ohjelmasta riippuen [Schmiegelow ym., 2014]. Tavoitepitoisuudella tarkoitetaan sellaista veren leukosyyttipitoisuutta, joka lääkannoksia säätämällä pyritään saavuttamaan. Esimerkiksi NOPHO-ALL 2008 -hoito-ohjelmassa tavoitteena oleva leukosyyttipitoisuus on ylläpitovaiheessa 1.5 – 3.0 solua/nl.⁵ Tässä hoito-ohjelmassa 6MP:n ja MTX:n alkuannokset suhteutetaan ylläpitovaiheessa potilaan kehon pinta-alaan (*body surface area*). Hoito aloitetaan 75 mg/m² 6MP-annoksella ja 20 mg/m² MTX-annoksella kaikissa ALL:n hoidon riskiryhmissä. Alkuannokseen voi kuitenkin vaikuttaa potilaan TPMT-genotyyppi, josta kerrotaan tarkemmin kappaleessa 2.3. Mikäli potilas on TPMT-genotyypiltään TPMT-heterosygootti tai TPMT-homosygootti, 6MP:n alkuannoksena käytetään 50 mg/m² tai 5 mg/m² annosta, tässä järjestyksessä. Hoidon aloittamisen jälkeen lääkannostuspäätöksissä seurataan leukosyytti- ja neutrofiilipitoisuuksia seuraavien suuntaa-antavien hoito-ohjeiden mukaisesti (Nielsen ym. [2017], liitemateriaali):

1. Jos leukosyyttipitoisuus on 1.0 – 1.5 solua/nl, 6MP- ja MTX-annokset on puolitettava.
2. Jos leukosyyttipitoisuus on alle 1.0 solua/nl, 6MP- ja MTX-lääkitys on hetkellisesti lopetettava. Hoitoa jatketaan 75 %:lla aiemmin annetusta lääkityksestä, kun leukosyyttipitoisuus on kasvanut suuremmaksi kuin 1.5 solua/nl.
3. Jos neutrofiilipitoisuus on pienempi kuin 0.5 solua/nl, 6MP- ja MTX-lääkityksen pienentämistä tai hetkellistä keskeyttämistä on harkittava.
4. Jos leukosyyttipitoisuus on yli 3.0 solua/nl yli kahden viikon lääkityksen jälkeen, 6MP- ja MTX-annoksia on kasvatettava noin 20 %.

Edellisten lisäksi hoito-ohjeissa tarkkaillaan mm. veren verihiutalepitoisuutta sekä maksa-arvoja, sillä 6MP ja MTX ovat maksalle myrkyllisiä aineita. Lääkannoksen muutoksia tehdään kuitenkin harvoin maksa-arvojen tai epänormaalin verihiutalepitoisuuden vuoksi. Tutkimusten mukaan riski pysyvälle tai merkittävälle maksavauriolle 6MP- ja MTX-lääkityksen vuoksi on pieni [Schmiegelow ym., 2014].

Lääkannosten pienennyksillä ja hoidon keskeytyksillä ylläpitovaiheessa on havaittu olevan yhteys ALL:n uusiutumiseen [Schmiegelow ym., 2014]. Toisaalta NOPHO-ALL

⁵ Vertailun vuoksi terveiden 2 – 6-vuotiaiden lasten veren leukosyyttipitoisuuden vaihteluväli on karkeasti ottaen 5 – 13 solua/nl [Synlab, 2017]. Neutrofiilien, jotka ovat leukosyyttien osajoukko, vastaava pitoisuuden vaihteluväli on noin 1.5 – 8.3 solua/nl.

1992 -hoito-ohjelmasta kertyneestä aineistosta havaittiin, että liian voimakkaan ja pitkäaikaisen 6MP-lääkityksen ja myöhemmin ilmenevien toisten syöpien kehittymisen välillä on yhteys [Schmiegelow ym., 2009]. Sopivan suuruisen lääkituksen löytäminen ja ylläpitäminen on siten keskeistä hoidon onnistumisen kannalta.

2.3 6MP:n yksilöllinen vaikutus leukosyyttipitoisuuteen

Lääkäri voi joutua tekemään hoito-ohjeista poikkeavia lääkeannostuspäätöksiä tavoitellessaan leukosyyttipitoisuuden tavoitepitoisuutta. Päätöksentekoa vaikeuttaa se, että 6MP:n vaikutus näkyy veren leukosyyttipitoisuuksissa viiveellä, usein vasta viikkojen kuluttua lääkkeen antamisesta [Gaynon, 2017]. Lisäksi lääkkeen vaikutus potilaaseen on yksilöllistä, sillä sen hyötyosuudessa on suurta vaihtelua yksilöiden välillä [Lafolie ym., 1986]. 6MP:n päästyä verenkiertoon sen muuntumiseen eri metaboliiteiksi vaikuttaa entsyymi tiopuriinimetyylitransferaasi (TPMT), jonka aktiivisuudessa on myös yksilöiden välistä vaihtelua. Yksilöillä, joilla on korkea TPMT-aktiivisuus, annetusta 6MP:stä muuntuu vähemmän sen aktiiviseksi metaboliitiksi, 6-tioguaaniiniksi (TGN) [Lennard ym., 1990]. TGN on ALL:n hoidon kannalta tärkeä metaboliitti, sillä 6MP:n solumyrkyllisyys perustuu pääosin TGN:n deoksimuotoon (TGN_{DNA}), joka liittyy potilaan DNA:han ja saa aikaan syöpäsolujen solukuolemia [Nielsen ym., 2017].

ALL-potilaat voidaan jaotella karkeasti pienenevän TPMT-aktiivisuuden mukaan kolmeen TPMT-genotyyppiin⁶, jotka ovat TPMT-villityyppi (86 – 97 % väestöstä), TPMT-heterotsygootti (3 – 14 % väestöstä) sekä TPMT-homotsygootti (alle 0.5 % väestöstä) [Nguyen ym., 2011]. TPMT-villityyppiä olevan potilaan tavallinen 6MP-annos voi olla TPMT-heterotsygooteille ja -homotsygooteille vaarallinen, sillä annoksen solumyrkyllinen vaikutus voi olla liian suuri TPMT-entsyymien ollessa vähemmän aktiivinen. Tämä pätee erityisesti TPMT-homotsygooteihin, joilla TPMT-entsyymi ei toimi lainkaan. Pelkkä TPMT-genotyyppi ei kuitenkaan määritä potilaalla ilmenevää TPMT-aktiivisuutta, sillä esimerkiksi Coulthard ym. [1998] mukaan väestön TPMT-aktiivisuuksissa esiintyy vaihtelua myös TPMT-genotyyppien sisällä.⁷

Alentuneen elimistön puolustuskyvyn vuoksi 6MP-hoidosta voi lisäksi aiheutua tulehduksia, jotka voivat olla vaarallisia [Lohi ym., 2013]. Lund ym. [2011] mukaan NOPHO-ALL 1992 ja NOPHO-ALL 2000 -hoito-ohjelmissa 72 % hoitoon liittyvistä kuolemista johtui tulehduksesta. Vaarallisuutensa lisäksi tulehdukset hankaloittavat leukosyyttipitoisuuden tavoitepitoisuuden saavuttamista, sillä tulehdusten aikana leukosyyttipitoisuudet voivat nousta tai laskea [Gaynon, 2017].

⁶ Genotyyppi tarkoittaa yksilön DNA:n sisältämien geenien muodostamaa kokonaisuutta.

⁷ Tässä tarkoitetaan TPMT-fenotyyppiin liittyvää mittausta eli mittausta siitä, minkälainen TPMT-aktiivisuus potilaalla *ilmenee* ympäristön ja TPMT-geenin vuorovaikutuksen vuoksi. Tekstissä mainitussa tutkimuksessa TPMT-aktiivisuuden mittausta oli tehty leukosyyteistä.

3 Aineisto

Tässä työssä käytettävä aineisto sisältää ALL:n ylläpitovaiheen hoidon aikana kerätyt tiedot 23 suomalaiselta alle 18-vuotiaalta potilaalta. Aineisto on peräisin Tampereen yliopistollisesta sairaalasta.

Potilaan hoidon aikana mitattujen muuttujien muodostamaa moniulotteista aikasarjaa kutsutaan tässä työssä potilaan hoitohistoriaksi. Hoitohistoriamuuttujat ovat annettuja lääkeannoksia lukuun ottamatta osajoukko täydellisessä verenkuvassa mitattavista muuttujista. Hoitohistorioiden lisäksi potilailta on mitattu eräät taustatietomuuttujat. Aineiston muuttujat on koottu taulukkoon 1.

Taulukko 1: Potilaiden hoitohistorioiden muuttujat ja taustatietomuuttujat aineistossa.

Hoitohistorian muuttuja	Yksikkö
Veren leukosyyttipitoisuus	solua/nl
Veren neutrofilipitoisuus	solua/nl
6MP-annos	mg/päivä
MTX-annos	mg/viikko
C-reaktiivisen proteiinin pitoisuus veressä	mg/l
Keskimääräinen punasolun tilavuus	fl
Alaniiniaminotransferaasi	yksikköä/nl
Hemoglobiini	g/l
Veren verihiutalepitoisuus	verihiutaletta/nl
Taustatietomuuttuja	Yksikkö
Potilaan pituus hoidon alussa	cm
Potilaan paino hoidon alussa	kg
TPMT-aktiivisuuteen liittyvä geeni	TPMT-genotyyppi

Potilaiden hoitohistorioiden mittausaika on yksi päivä ja pituus vaihtelee hieman yli 400 päivästä yli 500 päivään (minimi: 227, Q_1 : 442, Md : 506, Q_3 : 524, maksimi: 553). Työn kannalta kiinnostavimmat hoitohistoriamuuttujat ovat veren leukosyytti- ja neutrofilipitoisuudet, joiden mallintaminen on kiinnostuksen kohteena. Tässä työssä mallien vastemuuttujana käytetään leukosyyttipitoisuutta, sillä tiedossa on, että aineiston potilaita koskevista lääkeannostuspäätöksistä on seurattu pääasiassa potilaan veren leukosyyttipitoisuutta. Leukosyyttipitoisuuksista on hoitohistorioissa enemmän havaintoja kuin neutrofilipitoisuuksista ja jokaisen potilaan hoitohistoriassa leukosyyttipitoisuus on mitattu hoidon ensimmäisenä päivänä.

Potilaiden päivittäiset 6MP- ja MTX-annokset ovat tiedossa hoidon jokaiselta päivältä kaikissa hoitohistorioissa. Muissa hoitohistoriamuuttujissa on paljon puuttuvaa tietoa, sillä ne on mitattu tarpeen mukaan vain päivinä, joina potilas on käynyt lääkärissä. Tällaisia päiviä on useimmilla potilailla noin 40 (minimi: 20, Q_1 : 39, Md : 42, Q_3 : 46, maksimi: 102) ja ne ovat tyypillisesti 1 – 2 viikon välein. Osa potilaista on kuitenkin joutunut hoidon aikana sairaalahoitoon tulehduksen tai muun syyn vuoksi. Tämän seurauksena mittauksia on saatettu välillä tehdä päivittäin.

Aineiston taustatietomuuttujat sisältävät potilaiden pituudet ja painot hoidon alussa. Näiden lisäksi TPMT-genotyyppi on mitattu potilailta 1 – 10. Tieto TPMT-genotyypistä puuttuu muilta potilailta. Kymmenestä mitatusta potilaasta ainoastaan potilas 5 on TPMT-heterotsygootti, muut ovat TPMT-villityyppejä. Mitatuissa potilaissa ei siten ole yhtään TPMT-homotsygoottia.

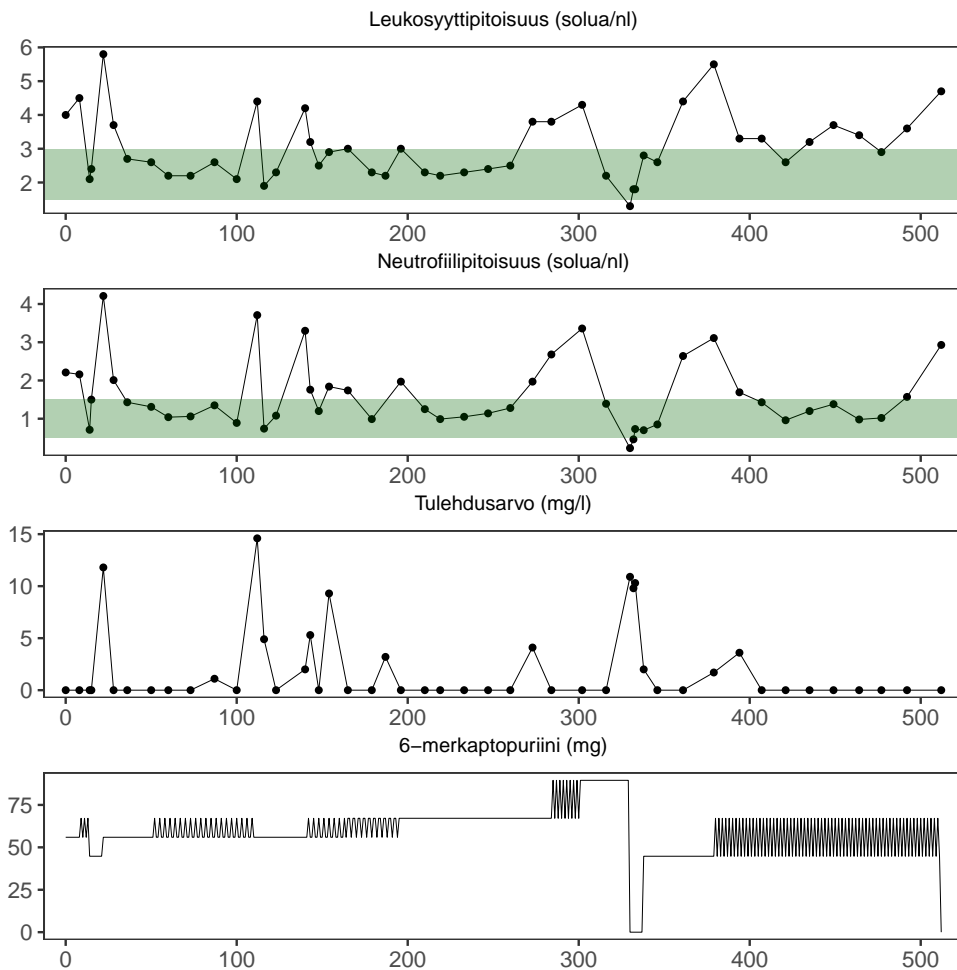
Tässä työssä leukosyyttipitoisuutta mallinnetaan pääasiassa 6MP-annoksen ja veren C-reaktiivisen proteiinin pitoisuuden suhteen. C-reaktiivisen proteiinin pitoisuus veressä kasvaa monenlaisten tulehdusten seurauksena [Eskelinen, 2016]. Bakteeritulehdukset kasvattavat pitoisuutta enemmän kuin virustulehdukset. Mittaukseen potilaan C-reaktiivisen proteiinin pitoisuudesta veressä viitataan tässä työssä jatkossa potilaan tulehdusarvolla. 6MP-annosten ja tulehdusarvon lisäksi myös potilaiden painoa ja pituutta hoidon alussa hyödynnetään kappaleessa 6 esiteltävässä mallissa. Mallinnuksen ulkopuolelle jätettyjen muuttujien hyödyntämistä myöhemmin kehitettävissä malleissa pohditaan kappaleessa 7.2.2.

Kuvaan 1 on visualisoitu erään potilaan hoitohistoria työn mallinnuksessa käytettävien hoitohistoriamuuttujien sekä neutrofilipitoisuuden osalta. Kuvan perusteella potilaan leukosyytti- ja neutrofilipitoisuuksien kehitys hoidon aikana näyttää samankaltaista. Hoito on onnistunut hyvin, sillä potilaan leukosyyttipitoisuudet on kyetty pitämään tavoitearvoissa suuren osan hoidon kestosta ja lääkitys on ollut pääosin tasaista.

Tulehdusarvon ja leukosyyttipitoisuuden välillä on kuvan perusteella yhteys. Esimerkiksi noin päivien 25 ja 100 kohdalla likipitäen tasaisen lääkityksen aikana havaitut korkeat tulehdusarvot näkyvät leukosyyttipitoisuuksissa selkeinä huippuina. Toisaalta esimerkiksi noin päivien 125 – 150 aikana havaittujen tulehdusarvohuippujen yhteys leukosyyttipitoisuuksiin ei vaikuta yhtä ilmeiseltä. Myöskään noin päivän 325 ja 400 aikaiset tulehdusarvohuiput eivät ole yhteydessä kasvaneisiin leukosyyttipitoisuuksiin samalla tavalla kuin hoidon alussa. Päivän 325 osalta tämän tosin selittänee samaan aikaan keskeytettyinä ollut lääkitys, joka johtunee korkealle nousseen tulehdusarvon ja matalalle pudonneen leukosyyttipitoisuuden yhdistelmästä.

Leukosyyttipitoisuuksien putoaminen tällä hoidon hetkellä saattaa johtua siitä, että potilaan lääkitys on kasvatettu potilaalle liian korkeaksi noin päivän 300 kohdalla. Lääkäri on keskeyttänyt lääkityksen noin päivän 325 kohdalla, jonka jälkeen leukosyyttipitoisuuden nähdään nousevan elimistön puolustautumisen kannalta turvallisemmalle tasolle. Tämän jälkeen lääkitystä on jatkettu pienemmällä annoksella. Noin päivänä 375 tehdyn mittauksen jälkeen lääkäri on jälleen päättänyt nostaa lääkitystä kenties siksi, että leukosyyttipitoisuus on noussut tavoitevälin yläpuolelle. Tällä kertaa lääkeannoksen nosto on kuitenkin tehty maltillisemmin, ja suurinpiirtein päivästä 400 lähtien leukosyyttipitoisuus onkin palannut tavoitevälin ylärajan tuntumaan. Hoitohistorian lopussa leukosyyttipitoisuus lähtee kuitenkin yllättäen nousuun. Kuvaa katsomalla tälle on hankala keksiä syytä. Mahdollinen selitys on, että potilas on jättänyt lääkkeen ottamatta tai potilaalla on ollut virustulehdus, joka ei ole näkynyt tulehdusarvossa, mutta on kasvattanut potilaan veren leukosyyttipitoisuutta.

Kokonaisuudessaan kuvasta voidaan todeta, että ajan suhteen tarkasteltuna lääkitys ja tulehdusarvo näyttävät selittävän leukosyyttipitoisuuden vaihtelua, mutta myös selittämätöntä vaihtelua jää jäljelle. Samankaltaisen päätelmän voi tehdä tarkastelemalla myös muita aineiston potilaita.



Kuva 1: Potilaan 1 hoitohistoria työn mallinnuksessa käytettävien muuttujien sekä neutrofiilipitoisuuden osalta. Vaaka-akselilla on hoitopäiviä. Havainnot on yhdistetty kolmessa ylimmässä aikasarjassa viivoin. Neutrofiili- ja leukosyyttipitoisuuden aikasarjoissa olevat vihreät alueet kuvaavat eräitä solupitoisuuksien tavoitevälejä: 1.5 – 3.0 solua/nl leukosyyteille ja 0.5 – 1.5 solua/nl neutrofileille. Sahalaitakuvio 6MP:n aikasarjassa osoittaa, että annosta on vaihdeltu vuoropäivin.

4 Kirjallisuudessa esitetyt mallit

Tässä työssä tehtävä mallinnus pohjautuu Friberg ym. [2002] ja Jayachandran ym. [2014, 2015] malleihin, jotka esitellään tässä kappaleessa. Mallien esittelyn lisäksi myös mallinnuksessa käytettyjä aineistoja ja mallinusoletuksia tarkastellaan.

4.1 Friberg ym. [2002] malli syöpälääkkeiden aiheuttamalle myelosupressiolle

Friberg ym. [2002] tavoitteena oli mallintaa useiden eri syöpälääkkeiden *myelosupressiivinen* vaikutus *leukopoiesiin*. Leukopoiesi tarkoittaa prosessia, jossa leukosyytit kehittyvät luuytimen kantasoluista. Myelosupressio tarkoittaa luuytimen solutuotannon heikkenemistä.

Mallin kehittämisen taustalla oli pyrkimys siihen, että samalla mallilla voisi mallintaa useiden eri syöpälääkkeiden myelosupressiivisen vaikutuksen. Mallinnuksessa käytetty aineisto sisälsi tiedot lääkekuurista ja leukosyytti- sekä neutrofilipitoisuuksista potilailta, joita oli kutakin lääkitty yhdellä kuudesta eri syöpälääkkeestä. Syöpälääkkeet olivat dose-takseli, paklitakseli, etoposidi, 2-deoksi-2-metyylidenesytidiini, irinotekaani ja vinfluniini. Friberg ym. [2002] jakoivat kehittämänsä mallin parametrit lääkekohtaisiin ja potilaskohtaisiin parametreihin. Mallin yleistyvyyttä usealle syöpälääkkeelle arvioitiin kiinnittämällä potilaskohtaiset parametrit alustavassa mallinnuksessa saatuihin tyypillisen potilaan parametrien arvoihin ja estimoimalla ainoastaan lääkekohtaiset parametrit kullekin syöpälääkkeelle erikseen. Näin estimoituja lääkekohtaisten parametrien arvoja tulkitsemalla artikkelin tekijät totesivat, että eri syöpälääkkeiden vaikutus leukopoiesiin on mahdollista kuvata yhdellä mallilla.

Artikkelissa esitetty malli on differentiaaliyhtälöryhmä, joka kuvaa leukopoiesia ajan funktiona:

$$\begin{aligned}\frac{dS}{dt} &= k_{prol}S [1 - E_{drug}(m)] \left(\frac{L_0}{L}\right)^\gamma - k_{tr}S \\ \frac{dC^{(1)}}{dt} &= k_{tr}S - k_{tr}C^{(1)} \\ \frac{dC^{(2)}}{dt} &= k_{tr}C^{(1)} - k_{tr}C^{(2)} \\ \frac{dC^{(3)}}{dt} &= k_{tr}C^{(2)} - k_{tr}C^{(3)} \\ \frac{dL}{dt} &= k_{tr}C^{(3)} - k_L L.\end{aligned}\tag{1}$$

Leukopoiesimallin (1) tilamuuttujien ja parametrien selitteet on koottu taulukkoon 2. Kukin mallin tilamuuttujista S , $C^{(1)}$, $C^{(2)}$, $C^{(3)}$ ja L kuvaa solupitoisuutta eräässä leukosyyttien erikoistumisvaiheessa. Erikoistumisvaiheet seuraavat toisiaan, sillä jokaisen erikoistumisvaiheen solupitoisuuden muutosnopeus on edellisessä ja nykyisessä erikoistumisvaiheessa olevan solupitoisuuden funktio. Tilamuuttuja S kuvaa erikoistumattomien kantasolujen pitoisuutta ja L kypsyneiden leukosyyttien pitoisuutta. Tilamuuttujat $C^{(1)}$, $C^{(2)}$ ja $C^{(3)}$ kuvaavat solupitoisuuksia näiden vaiheiden välille jäävissä erikoistumisvaiheissa, joita mallissa oletetaan olevan kolme. Erikoistumisvaiheet sisällytettiin malliin, jotta leu-

kopioeesissa oleva viive voitaisiin mallintaa. Mallin tilamuuttujista ainoastaan L havaitaan, sillä valmiiksi kehittyneet leukosyytit ovat verenkierrassa, josta ne voidaan mitata. Jäljelle jäävien tilamuuttujien arvoja ei havaita, koska näissä kehitysvaiheissa olevat solut kehittyvät luuytimessä, josta solupitoisuuksien mittauksia on vaikea tehdä.

Mallissa esiintyvät parametrit k_{prol} , k_{tr} ja k_L säätelevät kantasolujen jakautumisnopeutta, solujen erikoistumisnopeutta sekä leukosyyttien poistumisnopeutta verenkierrasta. Estimoitavien parametrien lukumäärän vähentämiseksi Friberg ym. [2002] kuitenkin olettivat, että $k_L = k_{prol} = k_{tr}$. Parametrin k_{tr} tulkinnan helpottamiseksi se asetettiin riippumaan parametrista t_{MTT} kaavalla

$$k_{tr} = \frac{4}{t_{MTT}}. \quad (2)$$

Parametri t_{MTT} (*MTT*, *mean transfer time*) kuvaa keskimääräistä aikaa, joka kuluu kantasolun kypsymiseen valmiiksi leukosyytiksi. Kaavassa (2) oleva luku 4 tulee mallin erikoistumisvaiheiden välisten siirtymien lukumäärästä.

Syöpälääke vaikuttaa mallissa kantasolupitoisuuden muutosnopeuteen syöpälääkkeen vaikutusta kuvaavan funktion $E_{drug}(m)$ kautta. Artikkelissa esiteltiin funktiolle kaksi erilaista muotoa:

$$E_{drug}(m) = Km \quad (3)$$

$$E_{drug}(m) = \frac{E_{max}m}{E_{C50} + m}, \quad (4)$$

jossa m kuvaa syöpälääkkeen vaikuttavan aineen konsentraatiota ja K , E_{max} sekä E_{C50} ovat parametreja. Kaksi parametria sisältävässä funktiossa (4) lääkkeen vaikutus voi olla funktiota (3) monimutkaisempi. Molemmat funktiot saavat arvon nolla, jos $m = 0$. Tällöin leukopoeesi toimii ilman syöpälääkkeen vaikutusta. Selvittääkseen syöpälääkkeiden vaikuttavien aineiden konsentraatioprofiilit aineistonsa potilaille, Friberg ym. [2002] hyödynsivät aineiston syöpälääkkeille kehitettyjä farmakokineettisiä malleja. Joidenkin lääkkeiden osalta konsentraatioprofiilit voitiin myös laskea aineistosta.

Kantasolupitoisuuden muutosnopeuteen vaikuttaa leukopoeesimallissa syöpälääkkeen vaikutuksen lisäksi termi

$$\left(\frac{L_0}{L}\right)^\gamma, \quad (5)$$

jota kutsutaan takaisinkytkennäksi. Takaisinkytkentä kiihdyttää kantasolujen tuotantoa suhteessa parametriin L_0 , joka kuvaa potilaan veren leukosyyttipitoisuuden tasapainopitoisuutta ilman lääkitystä. Jos veren leukosyyttipitoisuus on parametria pienempi, tuotanto kasvaa, ja jos pitoisuus on parametria suurempi, tuotanto pienenee. Kun $L = L_0$, takaisinkytkentä saa arvon 1. Parametri γ vaikuttaa takaisinkytkentäfunktion muotoon. Takaisinkytkennän sisällyttäminen leukopoeesimalliin tekee siitä epälineaarisen, jonka vuoksi differentiaaliyhtälöryhmän ratkaiseminen ei onnistu suljetussa muodossa, vaan ratkaisemisessa on turvauduttava numeerisiin menetelmiin.

Artikkelin malli sovitettiin epälineaarisenä sekamallina käyttäen vastemuuttujana sekä leukosyytti- että neutrofilipitoisuutta. Malli sovitettiin molempien vastemuuttujien tapauksessa kutakin syöpälääkettä kohden kummallakin syöpälääkkeen vaikutusta kuvaavalla funktiolla (3) ja (4). Taulukko 3 esittää tunnuslukuja parametrien γ , L_0 ja t_{MTT}

Taulukko 2: Friberg ym. [2002] esittelemän leukopioeesimallin tilamuuttujat ja parametrit selitteineen.

Tilamuuttuja	Selite	Yksikkö
S	Kantasolupitoisuus luuytimessä	solua / nl
$C^{(1)}$	Solupitoisuus 1. erikoistumisvaiheessa	solua / nl
$C^{(2)}$	Solupitoisuus 2. erikoistumisvaiheessa	solua / nl
$C^{(3)}$	Solupitoisuus 3. erikoistumisvaiheessa	solua / nl
L	Leukosyyttipitoisuus veressä	solua / nl
Parametri	Selite	Yksikkö
k_{prot}	Kantasolujen jakautumisnopeus	per päivä
t_{MTT}	Keskimääräinen leukosyytin kypsymiseen kuluva aika	päivää
k_L	Leukosyyttien poistumisnopeus verenkierrosta	per päivä
L_0	Leukosyyttien tasapainopitoisuus ilman lääkitystä	solua / nl
γ	Takaisinkytkennän muotoparametri	-
K	Syöpälääkkeen vaikutuksen muunnoskerroin	per konsentraatio*
E_{C50}	Syöpälääkkeen puolittaisen vaikutuksen konsentraatio	konsentraatio*
E_{max}	Syöpälääkkeen maksimivaikutus	-

*Syöpälääkkeiden konsentraatioiden yksikköä ei mainittu tarkasti.

Taulukko 3: Tunnuslukuja leukopioeesimallin (1) parametrien γ , L_0 ja t_{MTT} estimaateille yli malliver-sioiden, jossa leukosyyttipitoisuus oli vastemuuttujana. Parametriestimaatit ovat syöpälääkkeestä riippu-matta melko samanlaisia. Parametrit E_{max} , E_{C50} ja K jätettiin tämän taulukon ulkopuolelle, sille niiden estimaateissa oli eroja eri syöpälääkkeiden välillä.

Parametri	\bar{x}	min	Q_1	Md	Q_3	max
γ	0.17	0.12	0.15	0.17	0.18	0.24
L_0	7.30	6.74	7.10	7.21	7.50	8.10
t_{MTT}	4.93	3.73	4.67	5.17	5.50	5.62

estimaateista. Potilaiden välistä vaihtelua mallinnettiin sekamallissa log-normaalijakau-malla vain parametrien L_0 , t_{MTT} ja K (tai E_{C50}) osalta, sillä mallin muiden parametrien ei oletettu riippuvan potilaasta. Artikkelista ei käy ilmi, mitä alkuarvoja mallin tilamuut-tujille käytettiin.

4.2 Jayachandran ym. [2014, 2015] leukopioeesimalli ja 6MP:n farmakokineettinen malli

4.2.1 Leukopioeesimalli

Jayachandran ym. [2014] pääasiallisena tavoitteena oli kehittää malli veren leukosyytti-pitoisuuksien ennustamiseen lasten ALL:n 6MP-hoidon aikana. Leukosyyttipitoisuuksien ennustamisen taustalla oli pyrkimys helpottaa ALL:n hoitoon liittyvien lääkeannostuspää-tösten tekoa laskemalla mallin avulla potilaalle optimaalinen lääkitys. Mallista pyrittiin rakentamaan sellainen, että sen estimointi onnistuisi minimaalisella aineiston määrällä.

Malli on artikkelissa kokonaisuus, joka muodostuu alimallista leukopioeesille, 6MP:n farmakokineetikalle ja punasolujen keskimääräiselle koolle hoidon aikana. Tässä kappalees-sa keskitytään näistä ensimmäiseen, mutta viitataan myös artikkelissa esiteltyyn 6MP:n farmakokineettiseen malliin sen yksityiskohtia läpikäymättä. Lähes vastaava 6MP:n far-makokineettinen malli esiteltiin Jayachandran ym. [2015] artikkelissa, jota käsitellään kap-paleessa 4.2.2. Punasolujen keskimääräistä kokoa mallintava alimalli ei ole tämän työn kannalta kiinnostava.

Jayachandran ym. [2014] leukopoeiesimalli on mallin (1) innoittama. Se sisältää leukosyyttien erikoistumisvaiheiden solupitoisuuksia kuvaavat tilamuuttujat S , $C^{(1)}$, $C^{(2)}$, $C^{(3)}$ ja L , joista viimeistä neljää mallinnetaan samoilla differentiaaliyhtälöillä kuin leukopoeiesimallissa (1). Toisin kuin Friberg ym. [2002], Jayachandran ym. [2014] eivät kuitenkaan oleta, että solujen erikoistumisnopeus k_{tr} ja leukosyyttien poistumisnopeus k_L olisivat yhtä suuria.

Kantasolupitoisuuden muutosnopeutta mallinnetaan differentiaaliyhtälöllä:

$$\frac{dS}{dt} = k_{pl}(L)S - E_{drug}(m)S - k_{tr}S,$$

jossa

$$k_{pl}(L) = k_{pl}^{max} \frac{\rho^\gamma}{\rho^\gamma + L^\gamma} \quad (6)$$

kuvaa leukopoeiesin takaisinkytkentää. Takaisinkytkennän suuruus on rajoitettu maksimiarvoon k_{pl}^{max} . Takaisinkytkentäfunktioon (5) verrattuna takaisinkytkentäfunktio (6) sisältää sen voimakkuuteen vaikuttavan parametrin ρ , joka on kuitenkin tulkinnaltaan lähellä takaisinkytkentäfunktiossa (5) esiintyvää parametria L_0 . Jayachandran ym. [2014] leukopoeiesimallin parametrien tulkinta on muilta osin sama kuin leukopoeiesimallissa (1).

Jayachandran ym. [2014] mallintavat 6MP:n myelosuppressiivisen vaikutuksen syöpälääkkeen vaikutusta kuvaavan funktion (4) avulla. Artikkelissa oletetaan luuytimessä kehittyvien punasolujen säilyttävän niihin sitoutuneen 6MP:n solumyrkyllisen metaboliitin, TGN:n, siirtyessään verenkiertoon. Siten mittauksia punasoluihin sitoutuneesta TGN:stä (TGN_{RBC}) hyödynnetään artikkelissa luuytimessä kehittyvien solujen solumyrkyllisyyden asteen mallintamiseen. Kaavassa (4) konsentraatio m on TGN_{RBC} pikomooleissa $8 \cdot 10^8$ punasolua kohti. Artikkelissa esitelty 6MP:n farmakokineettinen malli mallintaa 6MP:n edellä mainitussa yksikössä oleviksi TGN_{RBC} - ja metyylimerkaptopuriinikonsentraatioiksi. Metyylimerkaptopuriinit (MeMP) ovat toisia 6MP:n metaboliitteja, jotka eivät artikkelin mallikokonaisuudessa ole yhteydessä luuytimen solumyrkyllisyyden asteeseen.

Artikkelin leukopoeiesimallin ja 6MP:n farmakokineettisen mallin estimointiin käytettiin 27 iältään 18-76-vuotiaan potilaan aineistoa, joka oli peräisin Hindorf ym. [2006] tutkimuksesta, jossa oli kaikkiaan mukana 60 IBD-potilasta⁸. Tutkimuksessa mukana olleita potilaita oli lääkitty 6MP:llä tai atsatiopriinilla⁹ (AZA). Potilaita oli seurattu 20 viikon ajan, jonka aikana heiltä oli mitattu leukosyyttipitoisuus ja TGN_{RBC} - sekä MeMP-konsentraatiot 12 kertaa. Mainittu 27 potilaan joukko edusti tutkimuksessa ryhmää, jonka hoito onnistui tutkimuksen hoitoprotokollan mukaisesti ja ilman hoidon sivuvaikutuksia. Taulukossa 4 esitetty hoitoprotokolla perustui potilaiden jaotteluun painon perusteella. Painoon perustuvasta hoitoprotokollasta huolimatta Jayachandran ym. [2014] olettivat 6MP:n farmakokineettisen mallinsa estimoinnissa, että kunkin potilaan lääkekuuri vastaa 70 kiloisen potilaan lääkekuuria, koska potilaskohtaisia painoja ei ilmoitettu Hindorf ym. [2006] artikkelissa.

⁸ IBD, *inflammatory bowel disease*. IBD-sairauksia ovat Crohnin tauti (*Crohn's disease*) ja haavainen paksusuolen tulehdus (*ulcerative colitis*). Tutkimuksessa mukana olevilla potilailla oli jompi kumpi edellä mainituista sairauksista.

⁹ AZA muuntuu elimistössä ei-entsymaattisin reaktioiden 6MP:ksi, ja milligrammoissa olevan AZA-annoksen saa karkeasti ottaen muunnettua vastaavaksi milligrammoissa olevaksi 6MP-annokseksi jakamalla sen luvulla 2.08 [Sandborn, 2001].

Taulukko 4: Hindorf ym. [2006] tutkimuksessa seurattu hoitoprotokolla.

<i>Hoitoprotokolla AZA:n tapauksessa</i>			
Paino	Viikko 1 (mg/pv)	Viikko 2 (mg/pv)	Viikot 3-20 (mg/pv)
45 – 55 kg	50	75	125
56 – 65 kg	50	100	150
66 – 75 kg	50	100	175
76 – 85 kg	50	125	200
>85 kg	75	150	225

<i>Hoitoprotokolla 6MP:n tapauksessa</i>			
Paino	Viikko 1 (mg/pv)	Viikko 2 (mg/pv)	Viikot 3-20 (mg/pv)
45 – 50 kg	25	25	50
51 – 70 kg	25	50	75
>70 kg	25	75	100

Hindorf ym. [2006] artikkelissa ei myöskään ilmoitettu potilaiden yksilöllisiä leukosyyttipitoisuuksia tai metaboliittikonsentraatioita, vaan saatavilla oli ainoastaan käytetyn 27 potilaan ryhmän metaboliittikonsentraatioista laskettu mediaani kullakin mittauskeralla. Näistä potilaista kahdeksan potilaan yksilölliset havainnot (leukosyyttipitoisuudet ja metaboliittikonsentraatiot) koko hoidon ajalta ovat kuitenkin esillä yksilöllistettyä lääkeannostusta käsittelevässä väitöskirjassa (Fransson [2007], s. 31 – 32). Jayachandran ym. [2014] keräsivät mallinnuksessa käyttämänsä aineiston näissä kahdessa lähteessä olevista kuvista.

Aineiston rakenteen vuoksi Jayachandran ym. [2014] tekivät mallinnuksen kahdessa vaiheessa sovittamalla aluksi nk. populaatiomallit erikseen leukopoeiesimallille ja 6MP:n farmakokineettiselle mallille. Populaatiomallien estimoinnissa aineistona käytettiin metaboliittikonsentraatioiden mediaaniaineistoa ja aineistoa, jossa kahdeksalta potilaalta saadut yksilölliset leukosyyttipitoisuudet oli keskiarvoistettu ajanhetkittäin. Koska käytössä olevan aineiston määrä oli vähäinen ja estimoitavia parametreja oli monta, Jayachandran ym. [2014] tekivät tämän jälkeen sensitiivisyysanalyysin, jolla populaatiomallien parametreista voitiin tunnistaa nk. sensitiiviset ja ei-sensitiiviset parametrit. Sensitiiviseksi parametriksi kutsutaan artikkelissa sellaista parametria, jolla on suuri vaikutus populaatiomallien sovitteiden arvoihin.

Populaatiomallinnuksen ja sensitiivisyysanalyysin jälkeen mallinnuksen toisessa vaiheessa aineiston kahdeksalle potilaalle, joiden yksilölliset havainnot olivat saatavilla, sovitettiin yksilölliset leukopoeiesimallit ja farmakokineettiset mallit. Yksilöllisiä malleja estimoidaessa ei-sensitiiviset parametrit kiinnitettiin populaatiomallinnuksessa saatuihin arvoihin. Edellä kuvatun menettelyn ideana on, että vain sensitiivisimmät parametrit estimoida potilaille saadaan yksilölliset mallit estimoida kaikkia parametreja potilas-kohtaisesti. Sensitiivisyysanalyysin tulokset ja kahdeksalle potilaalle estimoidut sensitiivisten parametrien arvot on esitetty taulukoissa 5 ja 6 leukopoeiesimallin osalta.

6MP:n farmakokineettinen malli ja leukopoeiesimalli sovitettiin erikseen, joten niiden yhdistelmä olettaa mallikohtaisten parametrien olevan toisistaan riippumattomia. Mallien sovittaminen tehtiin ratkaisemalla optimointiongelma, jossa minimoitiin painotettu neliöityjen virheiden summa havaitun aineiston ja mallien sovitteiden välillä. Artikkelista ei käy ilmi, mitä alkuarvoja leukopoeiesimallin tilamuuttujille käytettiin.

Taulukko 5: Jayachandran ym. [2014] sensitiivisyysanalyysin tulokset leukopoeesimallin osalta. Artikkelin tekijöiden mukaan parametrit ρ ja E_{C50} voidaan kiinnittää.

Parametri	Sensitiivisyys	Arvo populaatiomallinnuksessa	Yksikkö
k_{pl}^{max}	Sensitiivinen	0.3287	per päivä
γ	Sensitiivinen	0.4368	-
E_{max}	Sensitiivinen	0.0782	per päivä
k_{tr}	Sensitiivinen	0.1207	per päivä
k_L	Sensitiivinen	0.5346	per päivä
ρ	Ei-sensitiivinen	$8.2 \cdot 10^9$	solua/l
E_{C50}	Ei-sensitiivinen	84	pmol veressä*

*Kohti $8 \cdot 10^8$ punasolua.

Taulukko 6: Jaychandran ym. [2014] sensitiivisten parametrien estimaatit kahdeksalle potilaalle leukopoeesimallin osalta.

Potilas	k_{pl}^{max}	γ	E_{max}	k_{tr}	k_L
1	0.3380	0.4500	0.0600	0.1230	0.3653
2	0.3220	0.4650	0.0680	0.1200	0.6610
3	0.4340	0.3445	0.0810	0.1852	0.6637
4	0.3943	0.9154	0.2075	0.0926	0.5250
5	0.3920	0.1786	0.0930	0.1411	0.5145
6	0.3735	1.8000	0.2513	0.0858	0.5711
7	0.4470	0.9012	0.1664	0.1305	0.6790
8	0.8000	0.4818	0.2506	0.2663	1.2261

4.2.2 6MP:n farmakokineettinen malli

Jayachandran ym. [2015] farmakokineettinen malli kuvaa 6MP:n muuntumista TGN_{RBC} :ksi. Artikkelissa tehtävän mallinnuksen tavoitteena oli muodostaa populaatiojakauma farmakokineettisen mallin parametreille. Parametrien populaatiojakaumaa voisi tämän jälkeen hyödyntää priorijakaumana uusien potilaiden TGN_{RBC} -konsentraation ennustamisessa ja tätä kautta 6MP-hoidon optimoimisessa. Lähestymistapa mallinnukseen on artikkelissa bayesilainen.

Usein käytettyjen MCMC-menetelmien (esim. Gelfand ja Smith [1990]) sijaan Jayachandran ym. [2015] käyttivät mallinnuksessaan differentiaaliyhtälöryhmille soveltuvaa variaatio-Bayes-menetelmää [Láinez-Aguirre ym., 2016]. MCMC-menetelmiin verrattuna vähemmän laskennallisesti vaativassa variaatio-Bayes-menetelmässä mallin parametrien posteriorijakaumaa approksimoidaan todennäköisyysjakaumaperheellä. Artikkelissa approksimointiin käytettiin normaalijakaumaa.

Jayachandran ym. [2015] farmakokineettisen mallin 6MP:lle muodostaa differentiaaliyhtälöryhmä:

$$\begin{aligned}
 \frac{dX_{gut}}{dt} &= -k_{ab}X_{gut} + d(t) \\
 \frac{dX_{plasma}}{dt} &= k_{ab}X_{gut} - k_{el}X_{plasma} - \frac{k_{cm}X_{plasma}}{K + X_{plasma}} \\
 \frac{dX_{tgn}}{dt} &= \frac{v_{cm}k_{cm}X_{plasma}}{K + X_{plasma}} - k_{me}X_{tgn}.
 \end{aligned} \tag{7}$$

Mallin tilamuuttujien ja parametrien selitteet on koottu taulukkoon 7. Tulkinta mallille

Taulukko 7: Jayachandran ym. [2015] farmakokineettisen mallin tilamuuttujien ja parametrien selitteet.

Tilamuuttuja	Selite	Yksikkö
X_{gut}	6MP:n määrä vatsassa	pmol
X_{plasma}	6MP:n määrä veri-plasmassa	pmol
X_{tgn}	TGN _{RBC} -konsentraatio	pmol veressä*
*Kohti $8 \cdot 10^8$ punasolua.		
Parametri	Selite	Yksikkö
k_{ab}	6MP:n imeytymisnopeus	per päivä
k_{el}	6MP:n poistumisnopeus	per päivä
k_{cm}	6MP:n muuntumisnopeus TGN _{RBC} :ksi	pmol muuntunutta 6MP:tä päivässä
K	Michaelis-Menten vakio	pmol
k_{me}	TGN _{RBC} :n poistumisnopeus	per päivä

on seuraava: Hetkellä t annettu 6MP-annos $d(t)$ siirtyy aluksi potilaan vatsassa olevaksi 6MP:ksi X_{gut} , josta se imeytyy veri-plasmaan, X_{plasma} , nopeudella k_{ab} . Veri-plasmassa oleva 6MP vähenee kahdella tapaa: parametrin k_{el} määrittämä määrä poistuu kokonaan, ja ns. Michaelis-Menten-kinetiikkaan (Michaelis ja Menten [1913]) perustuvan termin

$$\frac{k_{cm}X_{plasma}}{K + X_{plasma}}$$

määrittämä määrä muuntuu TGN_{RBC}:ksi. Syntynyt TGN_{RBC}-konsentraatio pienenee nopeudella k_{me} .

Artikkelista ei käy ilmi, mitä alkuarvoja mallin tilamuuttujille käytettiin. Myöskään tilamuuttujan X_{tgn} differentiaaliyhtälössä olevan yksikkömuunnosvakion v_{cm} arvoa ei mainita. Lähes vastaavassa Jayachandran ym. [2014] farmakokineettisessä mallissa vastaavan yksikkömuunnosvakion arvona oli 1.

Jayachandran ym. [2015] käyttivät mallin parametrien estimointiin kahdesta eri tutkimuksesta kerättyä aineistoa. Ensimmäinen aineistoista oli kappaleessa 4.2.1 mainittu Franssonin [2007] väitöskirjasta saatu 8 potilaan aineisto. Toinen aineisto sisälsi TGN_{RBC}-konsentraatiot 23 potilaalta, joita oli lääkitty kahdeksan viikon ajan päivittäisellä 50 milligramman 6MP-annoksella. Aineisto oli peräisin Derijks ym. [2004] tutkimuksesta, jossa tutkittiin Crohnin tautia tai haavaista paksusuolen tulehdusta sairastavien potilaiden hoitoa 6MP:llä. Potilaat olivat 19-68-vuotiaita ja heitä oli tutkimuksessa kaikkiaan 30. Tutkimuksessa TGN_{RBC}-konsentraatio mitattiin lähtökohtaisesti kultakin potilaalta viisi kertaa, mutta joissakin tapauksissa hoito keskeytettiin sivuvaikutusten vuoksi.

Jayachandran ym. [2015] muodostivat mallilleen uskottavuusfunktion

$$L(y_i, D_i | \theta, \sigma_{tgn}) = \prod_{k=1}^{n_i} \left[\frac{1}{\sqrt{2\pi\sigma_{tgn}^2}} \exp \left(-\frac{(y_{ik} - \hat{y}_i(t_{ik}, \theta, D_i))^2}{2\sigma_{tgn}^2} \right) \right], \quad (8)$$

jossa D_i viittaa potilaan i saamiin 6MP-annoksiin. Uskottavuusfunktio on saatu olettamalla, että potilaan i TGN_{RBC}-konsentraation havainto y_{ik} on normaalijakautunut odotusarvolla $\hat{y}_i(t_{ik}, \theta, D_i)$ ja varianssilla σ_{tgn}^2 . Edellä $\hat{y}_i(t_{ik}, \theta, D_i)$ on mallin tilamuuttujan X_{tgn} ratkaisu potilaalle i hetkellä t_{ik} , jolloin havainto y_{ik} tehtiin. Parametrivektori θ koostuu taulukossa 7 olevista parametreista. Muuttuja n_i kuvaa potilaasta i olevien TGN_{RBC}-mittausten lukumäärää.

Posteriorijakauma potilaalle i on Bayesin kaavan perusteella:

$$p_i(\theta, \sigma_{tgn}|y_i, D_i) = \frac{L(y_i, D_i|\theta, \sigma_{tgn})p(\theta, \sigma_{tgn})}{p(y_i, D_i)}. \quad (9)$$

Jayachandran ym. [2015] muodostivat priorijakauman $p(\theta, \sigma_{tgn})$ etsimällä ensin suurimman uskottavuuden estimaatit parametrivektorille θ ja parametrille σ_{tgn} kunkin potilaan aineistossa. Näin saatuihin potilaskohtaisiin parametriestimaatteihin sovitettiin tämän jälkeen moniulotteinen log-normaalijakauma, jota käytettiin kaavassa (9) priorijakaumana $p(\theta, \sigma_{tgn})$. Muodostetun priorijakauman avulla kunkin potilaan posteriorijakauma $p_i(\theta, \sigma_{tgn}|y_i, D_i)$ ratkaistiin variaatio-Bayes-menetelmää käyttäen. Mallin tuntemattomille parametreille muodostettiin lopuksi populaatiojakauma simuloimalla ensin kunkin potilaan posteriorijakaumasta 10000 näytettä ja painottamalla näin saatuja otoksia potilaskohtaisten havaintojen lukumäärällä, joka vaihteli useasta lähteestä koostetun aineiston vuoksi. Lopullinen populaatiojakauma muodostettiin yhdistämällä painotetut otokset ja approksimoimalla saatua sekoitusjakaumaa moniulotteisella normaalijakaumalla. Edellä kuvatulla tavalla estimoidun populaatiojakauman parametrien likimääräiset marginaalijakaumat on esitetty taulukossa 8.

Jayachandran ym. [2015] pienensivät edelleen mallin estimoitavien parametrien lukumäärää tekemällä sensitiivisyysanalyysin kuten Jayachandran ym. [2014]. Sensitiivisyysanalyysin perusteella parametreilla k_{cm} ja k_{me} on suurin vaikutus mallin sovitteisiin. Parametrien k_{cm} ja k_{me} välillä havaittiin kuitenkin korkea korrelaatio, 0.96. Korrelaatio ei poistunut myöskään tilanteessa, jossa pelkästään k_{cm} ja k_{me} estimoitiin, ja muiden parametrien arvot asetettiin marginaalijakaumiensa keskiarvoihin. Jayachandran ym. [2015] mukaan sensitiivisyysanalyysin tulosten ja parametrien k_{cm} ja k_{me} välisen korkean korrelaation perusteella mallin parametreista riittää estimoida pelkästään parametri k_{cm} , ja muut parametrien arvot voidaan kiinnittää parametrien marginaalijakaumien keskiarvoihin. Estimoidussa populaatiojakaumassa muiden parametrien kuin parametrien k_{cm} ja k_{me} väliset korrelaatiot eivät olleet suuria.

Taulukko 8: Jayachandran ym. [2015] 6MP:n farmakokineettisen mallin parametrien marginaalijakaumat. Jakaumat ovat likimääräisiä, sillä ne esiteltiin artikkelissa vain graafisesti.

Parametri	Jakauma
$\log(k_{cm})$	$N(4.50, 1.20)$
$\log(k_{ab})$	$N(1.25, 1.39)$
$\log(k_{el})$	$N(1.11, 1.30)$
$\log(K)$	$N(2.50, 1.51)$
$\log(k_{me})$	$N(-2.40, 0.80)$
$\log(\sigma_{tgn})$	$N(0.60, 0.17)$

5 Bayesilainen differentiaaliyhtälömalli

Tässä kappaleessa esitellään työn ensimmäinen malli ja siihen liittyvät mallinnustulokset. Myös mallin parametrien estimointia sekä estimoinnin diagnostiikkaa tarkastellaan. Kappaleessa tehtävän mallinnuksen tarkoituksena on soveltaa edellä esiteltyjä malleja ja mallinnustuloksia käytössä olevaan potilasaineistoon.

5.1 Mallin kuvaus

Bayesilaisella differentiaaliyhtälömallilla viitataan jatkossa tässä kappaleessa esiteltävään mallikokonaisuuteen, joka muodostuu 6MP:n farmakokineettisen mallin (7) ja kappaleessa 4.2.1 kuvatun leukoioeesimallin yhdistelmästä. Kuva 2 esittää bayesilaisen differentiaaliyhtälömallin rakenteen kaaviona. Mallin tilayhtälö muodostuu differentiaaliyhtälöryhmästä

$$\begin{aligned}
 \frac{dX_{gut}}{dt} &= -k_{ab}X_{gut} + d(t) \\
 \frac{dX_{plasma}}{dt} &= k_{ab}X_{gut} - k_{el}X_{plasma} - \frac{k_{cm}X_{plasma}}{K + X_{plasma}} \\
 \frac{dX_{tgn}}{dt} &= \frac{v_{cm}k_{cm}X_{plasma}}{K + X_{plasma}} - k_{me}X_{tgn} \\
 \frac{dS}{dt} &= k_{pl}(L) - \frac{E_{max}X_{tgn}}{EC_{50} + X_{tgn}} - k_{tr} \\
 \frac{dC^{(1)}}{dt} &= k_{tr} \exp(S - C^{(1)}) - k_{tr} \\
 \frac{dC^{(2)}}{dt} &= k_{tr} \exp(C^{(1)} - C^{(2)}) - k_{tr} \\
 \frac{dC^{(3)}}{dt} &= k_{tr} \exp(C^{(2)} - C^{(3)}) - k_{tr} \\
 \frac{dL}{dt} &= k_{tr} \exp(C^{(3)} - L) - k_L,
 \end{aligned} \tag{10}$$

jossa

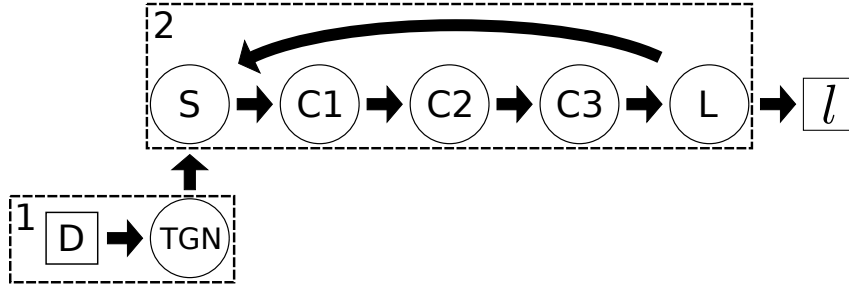
$$k_{pl}(L) = k_{pl}^{max} \frac{\rho^\gamma}{\rho^\gamma + \exp(L)^\gamma}.$$

Tilamuuttujia $S - L$ koskevat differentiaaliyhtälöt on edellä muunnettu logaritmisskaalaan ja ne kuvaavat siten muutosnopeuksia tilamuuttujien logaritmeissa. Muuntaminen on tehty siis esimerkiksi tilamuuttujan L tapauksessa laskemalla

$$\frac{d \log(L)}{dt} = \frac{d \log(L)}{dL} \frac{dL}{dt} = \frac{1}{L} \frac{dL}{dt} = \frac{1}{L} [k_{tr}C^{(3)} - k_L L] = k_{tr} \frac{C^{(3)}}{L} - k_L,$$

ja siirtymällä merkintään $\log(L) = L$ tilayhtälössä (10).

Kappaleeseen 3 viitaten käytössä oleva aineisto ei sisällä TGN_{RBC}-mittauksia, kuten Jayachandran ym. [2015] käyttämä aineisto. Tästä syystä farmakokineettisen mallin vastemuuttujana oletetaan olevan TGN-konsentraatio, ottamatta tarkemmin kantaa siihen, onko kyseessä TGN_{RBC} vai esimerkiksi kappaleessa 2.3 mainittu TGN_{DNA}. TGN-mittausten puuttuminen saa aikaan mallinnushaasteen, sillä tilayhtälön (10) sisältämää



Kuva 2: Bayesilaisen differentiaaliyhtälömallin rakenne. Mallin tilayhtälö koostuu 6MP:n farmakokineettisestä mallista (1) ja leukopoeiesimallista (2). Farmakokineettinen malli muuntaa potilaalle annetun lääkeannoksen, D , TGN:ksi. TGN vaikuttaa leukopoeiesiin, josta havaitaan leukosyyttipitoisuus l . Kaareva pitkä nuoli kuvaa leukopoeiesin takaisinkytkentää.

farmakokineettista mallia ei voida mittausten puuttuessa sovittaa ja tutkia erillään tilayhtälön leukopoeiesimallista.

6MP-annoksen yksikkönä oli farmakokineettisessä mallissa (7) pikomooli. Milligrammoista pikomooliksi muunnetun¹⁰ 6MP-annoksen käyttäminen tilayhtälössä (10) ei kuitenkaan tuottanut tämän kappaleen mallia varten tehdyssä alustavassa mallinnuksessa tyydyttävää farmakokineettista mallia 6MP:lle, sillä farmakokineettisen mallin avulla simuloidut TGN-konsentraatiot eivät pikomoolissa olevaa lääkeannosta käyttäen reagoineet lainkaan potilaan lääkityksen muutoksiin. Jos muunnosta pikomooliksi ei tehty, simuloidut TGN-konsentraatiot reagoivat paremmin lääkityksen muutoksiin. Tämän vuoksi lääkeannos $d(t)$ on tilayhtälössä (10) milligrammoissa farmakokineettisestä mallista (7) poiketen. Muilta osin tilayhtälössä esiintyvien tilamuuttujien ja parametrien tulkinta on kuten kappaleessa 4.

Kappaleessa 4.2 esitellyistä lähteistä ei selvinnyt, mitä alkuarvoja 6MP:n farmakokineettisen mallin ja leukopoeiesimallin tilamuuttujille oli käytetty. Estimoitavien parametrien lukumäärän minimoimiseksi leukopoeiesimallin alkuarvot tilayhtälössä (10) asetetaan siten, että ylläpitovaiheen hoidon alkaessa potilaan leukopoeiesin oletetaan olevan lähellä tasapainotilaa. Leukopoeiesi on tasapainotilassa, kun leukopoeiesimallin tilamuuttujien muutosnopeudet ovat nollia. Tämä saavutetaan tilamuuttujaa S lukuun ottamatta asetamalla

$$\begin{aligned}
 L(0) &= L_{init} \\
 C^{(3)}(0) &= \log(k_L) - \log(k_{tr}) + L(0) \\
 C^{(2)}(0) &= C^{(3)}(0) \\
 C^{(1)}(0) &= C^{(2)}(0) \\
 S(0) &= C^{(1)}(0).
 \end{aligned}$$

Edellä L_{init} on tilamuuttujan L alkuarvoa kuvaava parametri. Tilamuuttujan S muutosnopeus hoidon alkaessa riippuu tilayhtälössä (10) esiintyvien parametrien sekä parametrien L_{init} arvoista. Tilamuuttujien X_{gut} , X_{plasma} ja X_{tgn} alkuarvot oletetaan nolliksi. Siten 6MP ei vaikuta leukopoeiesiin hoidon alussa.

Bayesilaisen differentiaaliyhtälömallin uskottavuusfunktio muodostetaan kuten uskottavuusfunktio (8). Funktio on siten potilaalle i muotoa

¹⁰ Muunnos voidaan tehdä 6MP:n moolimassaa, 152.175 g/mol, hyödyntämällä [PubChem, 2018].

$$L(l_i, D_i | \theta, \sigma_{leuk}) = \prod_{k=1}^{n_i} \left[\frac{1}{\sqrt{2\pi\sigma_{leuk}^2}} \exp \left(-\frac{(l_{ik} - \hat{L}_i(t_{ik}, \theta, D_i))^2}{2\sigma_{leuk}^2} \right) \right]. \quad (11)$$

Edellä potilaan logaritmoitu leukosyyttipitoisuus l_{ik} on mallinnettu normaalijakautuneeksi tilamuuttujan L ratkaisun $\hat{L}_i(t_{ik}, \theta, D_i)$ ympärille keskihajonnalla σ_{leuk} . Parametrivektori θ muodostuu tilayhtälössä (10) esiintyvistä parametreista sekä parametrilla L_{init} . Uskottavuusfunktion perusteella malli voidaan nähdä regressiomallina, jonka sovite on tilayhtälön (10) tilamuuttujan L ratkaisu. Hoitohistorioissa olevan puuttuvuuden oletetaan olevan riippumaton tilamuuttujista ja parametreista. Malli sovitetaan kullekin potilaalle erikseen.

5.2 Parametrien priorijakaumat ja kiinnitetyt arvot

Bayesilaisen differentiaaliyhtälömallin estimoitavien parametrien priorijakaumat ja kiinnitettyjen parametrien arvot on koottu taulukkoon 9. Yksikkömuunnosvakion v_{cm} oletetaan olevan arvoltaan 1. Parametrit k_{ab} , k_{el} , K ja k_{me} kiinnitettiin kappaleessa 4.2.2 mainittuun sensitiivisyysanalyysin perusteella taulukon 8 mukaisten jakaumien odotusarvoihin. Samassa taulukossa mukana oleva parametri σ_{tgn} jätettiin mallin ulkopuolelle TGN_{RBC}-mittausten puuttumisen vuoksi. Parametrien E_{max} , k_{pl}^{max} , k_{tr} ja γ arvot asetettiin taulukon 5 mukaisiin parametristimaatteihin. Parametrin E_{C50} arvo asetettiin pienempään arvoon kuin taulukossa 5, sillä taulukon arvo oli estimoitu ehdolla, että 6MP:n yksikkönä on pikomooli.

Mallin estimoitaviksi parametreiksi valittiin parametrit k_{cm} , k_L , σ_{leuk} , ρ ja L_{init} , sillä malliversiossa, jossa estimoitaviksi parametreiksi valittiin kappaleessa 4.2 mainittujen sensitiivisyysanalyysien sensitiiviset parametrit, oli identifioituvuusongelmia. Mallin estimoitavien parametrien valinnassa hyödynnettiin MAP-estimointia (*maximum a posteriori*) ja lyhyitä MCMC-ketjuja, joiden avulla mallin sopivuutta aineistoon tarkasteltiin ennen lopullisten MCMC-ketjujen simulointia, jota käsitellään kappaleissa 5.3 ja 5.4. Edellä mainittujen parametrien valinta tuotti testatuista malliversioista parhaan tuloksen parametrien identifioituvuuden ja mallin aineistoon sopivuuden mielessä.

Estimoitaviksi valittujen parametrien priorijakaumina käytettiin logaritmiskaalassa olevia normaalijakaumia, jotka vastaavat log-normaalijakaumia lineaariskaalassa. Para-

Taulukko 9: Bayesilaisen differentiaaliyhtälömallin parametrien priorijakaumat ja kiinnitetyt arvot. Priorijakaumat ja arvot perustuvat Jayachandran ym. [2014, 2015] mallinnustuloksiin. Vasemmalla 6MP:n farmakokinetiikkaan vaikuttavat parametrit ja oikealla leukopoiesiin vaikuttavat parametrit. Parametrin L_{init} priorijakauman odotusarvo määräytyy potilaan ensimmäisen leukosyyttimittauksen perusteella.

Parametri	Priorijakauma tai arvo	Parametri	Priorijakauma tai arvo
$\log(k_{cm})$	$N(4.50, 1.80)$	$\log(k_L)$	$N(-0.4854, 0.4833)$
k_{ab}	3.4903	$\log(\sigma_{leuk})$	$N(0, 1000)$
k_{el}	3.0377	$\log(\rho)$	$N(\log(8.2), 1000)$
K	12.1825	L_{init}	$N(l_{i1}, 1)$
k_{me}	11.0232	E_{max}	0.0782
		γ	0.4368
		k_{tr}	0.1207
		k_{pl}^{max}	0.3287
		E_{C50}	9.00

metrin k_{cm} priorijakaumaksi asetettiin taulukon 8 mukainen marginaaliposteriorijakauma. Parametrin k_L priorijakauma saatiin sovittamalla log-normaalijakauma taulukon 6 mukaisiin potilaskohtaisiin parametriestimaatteihin. Edellä mainittujen jakaumien keskijajontoja kasvatettiin kertoimella 1.5, jotta priorijakaumista tulisi vähemmän informatiivisia. Parametreille σ_{leuk} ja ρ asetettiin epäinformatiiviset priorijakaumat. Parametrin ρ priorijakauman odotusarvo asetettiin taulukon 9 perusteella. Odotusarvo eroaa taulukossa olevasta arvosta kertoimella 10^{-9} , sillä Jayachandran ym. [2014] aineistossa leukosyyttipitoisuudet olivat kohti litraa verta nanolitrojen sijaan. Alkuarvoparametrin L_{init} priorijakauma asetettiin kunkin potilaan ensimmäisen leukosyyttimittauksen perusteella.

5.3 Estimointi

Estimoitaviksi valittujen parametrien posteriorijakauman simulointi tehtiin käyttämällä adaptiivista Metropolis-algoritmia, jonka ehdotusjakaumana on moniulotteinen normaali-jakauma [Vihola, 2012]. Adaptiivinen Metropolis-algoritmi on kuvattu algoritmossa 1.

Algoritmi 1 Adaptiivinen Metropolis-algoritmi

- 1: Valitse alkuarvot $\theta^{(0)}$ ja S_0 .
 - 2: *Toista iteraatiota $n = 1, \dots, N$ kohden:*
 - 3: Simuloi ehdotus $\theta^* = \theta^{(n-1)} + S_{n-1}U_n$, jossa $U_n \sim N(0, I)$.
 - 4: Laske $\alpha_n = \min \left\{ 1, \frac{f(\theta^*)}{f(\theta^{(n-1)})} \right\}$, jossa f on normeeraamaton posteriorijakauma.
 - 5: Todennäköisyydellä α_n , aseta $\theta^{(n)} = \theta^*$, muutoin $\theta^{(n)} = \theta^{(n-1)}$.
 - 6: Laske S_n kaavan (12) avulla.
-

Algoritmi koostuu tavallisesta Metropolis-askeleesta ja tämän jälkeen vaiheessa 6 tapahtuvasta ehdotusjakauman adaptioimisesta. Ehdotusjakauman adaptointi tarkoittaa sitä, että ehdotusjakauman kovarianssimatriisia muunnetaan simuloinnin aikana kohteena olevan posteriorijakauman näytteistämisen kannalta suotuisammaksi. Adaptoinnissa laskeaan uusi Cholesky-hajotelma ehdotusjakauman kovarianssimatriisille kaavan

$$S_n S_n' = S_{n-1} \left(I + \eta_n (\alpha_n - \alpha_*) \frac{U_n U_n'}{\|U_n\|^2} \right) S_{n-1}' \quad (12)$$

perusteella. Edellä S_n' on Cholesky-hajotelman S_n transpoosi, I on yksikkömatriisi, α_* on tavoitteena oleva hyväksymistodennäköisyys ja $\{\eta_n\}_{n \geq 1} \subset (0, 1]$ on nolaa lähestyvä, askelkokoa kuvaava lukujono. Lukujonon alkioiden ollessa lähellä nolaa, ehdotusjakauman kovarianssimatriisi pysyy lähes samana. Tässä työssä lukujono $\{\eta_n\}_{n \geq 1}$ on valittu siten, että sen järjestyslukua n vastaava alkio saadaan kaavasta $\eta_n = \min\{1, dn^{-2/3}\}$. Edellä d on estimoitavien parametrien lukumäärä, joka on bayesilaisen differentiaaliyhtälömallin tapauksessa 5. Tavoitteena olevan hyväksymistodennäköisyyden α_* arvona käytetään tässä työssä arvoa 0.234, sillä arvon voidaan eräin ehdoin osoittaa olevan optimaalinen tavoitteena oleva hyväksymistodennäköisyys [Roberts ym., 1997].

Algoritmin 1 vaiheessa 4 tarvittava normeeraamattoman posteriorijakauman arvo saadaan kaavasta

$$L(l_i, D_i | \theta, \sigma_{leuk}) p(\theta, \sigma_{leuk}),$$

jossa $L(l_i, D_i | \theta, \sigma_{leuk})$ on uskottavuusfunktio (11). Ratkaisemalla tilayhtälö (10) saadaan potilaasta i koskeva tilamuuttujan L ratkaisu $\hat{L}_i(t_{ik}, \theta, D_i), k = 1, \dots, n_i$. Sijoittamalla ratkaisu uskottavuusfunktioon (11) sen arvo saadaan laskettua. Priorijakauman $p(\theta, \sigma_{leuk})$ tiheysfunktion arvo lasketaan tulona taulukossa 9 esitettyjen jakaumien tiheysfunktioiden arvoista, joten parametrien priorijakaumat ovat bayesilaisessa differentiaaliyhtälömallissa toisistaan riippumattomia.

Algoritmin 1 toteutus muokattiin R-kielen [R Core Team, 2017] *bssm*-kirjaston [Helsinki ja Vihola, 2018] lähdekoodista sopivaksi. Tilayhtälön (10) ratkaisemisessa käytettiin C++-kielen *odeint*-kirjaston differentiaaliyhtälöryhmien numeerista ratkaisijaa [Ahnert ja Mulansky, 2011].

5.4 Konvergenssi

Bayesilainen differentiaaliyhtälömalli sovitettiin simuloimalla kunkin potilaan parametrien posteriorijakaumaa kolmella 100000 iteraation MCMC-ketjulla. Ketjujen alkuarvot valittiin satunnaisesti moniulotteisesta normaalijakaumasta, jonka komponenttijakaumien varianssit olivat suuria. Parametrin L_{imit} alkuarvo asetettiin kuitenkin potilaan ensimmäisen leukosyyttimittauksen logaritmiin kaikissa MCMC-ketjuissa. Ketjujen lämmitykseen käytettiin 25000 iteraatiota. Ehdotusjakauman kovarianssimatriisin adaptointia jatkettiin koko simuloinnin ajan. Ketjujen simulointi kesti rinnan ajettuna noin kolme vuorokautta 4-ytimisellä 2.2 gigahertsin Intel i7-3632QM -prosessorilla.

Tässä työssä MCMC-ketjujen konvergenssin tarkastelu tehdään tarkastelemalla ketjuja graafisesti, laskemalla efektiiviset otoskoot (*effective sample size*) sekä tunnusluvun \hat{R} (*potential scale reduction factor*) 95 %:n luottamusvälin yläraja potilaittain ja parametreittain. Tunnusluku \hat{R} mittaa simuloitujen MCMC-ketjujen stationaarisuutta ja sekoitumista vertailemalla ketjujen sisäistä ja välistä vaihtelua parametreittain (Gelman ym. [2014], s. 284 – 286). Gelman ym. [2014] (s. 287 – 288) mukaan arvoa 1.1 pienemmän \hat{R} -arvon saanutta MCMC-ketjua voidaan pitää lähtökohtaisesti konvergoineena tasapainojakaumaansa eli kiinnostuksen kohteena olevaan posteriorijakaumaan. Arvon 1.1 alittava \hat{R} ei takaa ketjun konvergenssia, mutta voi olla käytännöllisesti katsoen hyvä lähtökohta konvergenssin tarkasteluun. Tunnusluvun käyttö kuitenkin olettaa, että MCMC-ketjun stationaarinen jakauma on lähellä normaalijakaumaa (Gelman ym. [2014], s. 287). Siten tunnuslukua ei voi käyttää konvergenssikriteerinä esimerkiksi monihuippuisten posteriorijakaumien tapauksessa. Tunnusluvulle on mahdollista laskea luottamusvälin yläraja \hat{R}_u , joka perustuu sen otantajakauman approksimoimiseen F-jakaumalla (Gelman ja Rubin [1992], s. 465). Luottamusvälin ylärajan käyttäminen konvergenssista kertovana lukuarvona on tunnusluvun piste-estimaattia tiukempi ehto konvergenssille.

Efektiivinen otoskoko antaa arvion siitä, kuinka monen riippumattoman havainnon perusteella parametrin jakauman voidaan ajatella olevan estimoitu (Gelman ym. [2014], s. 286 – 287). MCMC-menettelmien tuottamat ketjut sisältävät autokorrelaatiota, jonka efektiivinen otoskoko huomioi. Gelman ym. suosittelevat simuloimaan niin pitkät MCMC-ketjut, että efektiivinen otoskoko on vähintään 10 kertainen simuloitujen ketjujen lukumäärään nähden.

Bayesilaisen differentiaaliyhtälömallin potilaskohtaiset posteriorijakaumat olivat potilaasta riippuen yksi- tai kaksihuippuisia. Kaksihuippuiset posteriorijakaumat havaittiin graafisen tarkastelun perusteella potilailla 10, 14, 15, 20 ja 21, joilla parametrikohdaiset ylärajan \hat{R}_u arvot olivat reilusti arvoa 1.1 suurempia. Näiden potilaiden MCMC-ketjujen

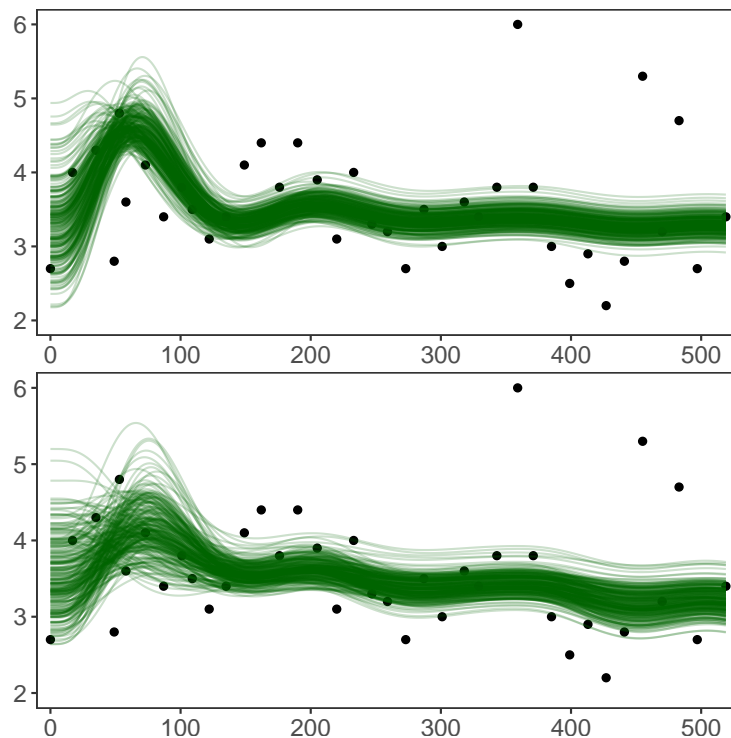
ei käytetyn simulointialgoritmin vuoksi oleteta edustavan hyvää otosta potilaiden posteriorijakaumista. Posteriorijakauman kaksihuippuisuus näkyi erityisen hyvin potilaan 15 MCMC-ketjuista, joista yksi oli simuloinut näytteitä pelkästään posteriorijakauman ensimmäisen huipun ympäristöstä, ja kaksi muuta pelkästään toisen huipun ympäristöstä. Muilla potilailla, joilla posteriorijakauma oli kaksihuippuinen, MCMC-ketjut siirtyivät huipusta toiseen ketjujen simuloinnin aikana.

Muiden kuin edellä mainittujen potilaiden posteriorijakaumat olivat MCMC-ketjujen perusteella yksihuippuisia. Potilaiden 9 ja 16 parametrisoitteiset ylärajan \hat{R}_u arvot ylittivät kuitenkin joidenkin parametrien tapauksessa arvon 1.1, joten näiden potilaiden MCMC-ketjujen ei katsottu konvergoineen potilaiden posteriorijakaumiin. Jäljelle jäävillä potilailla ylärajat \hat{R}_u alittivat arvon 1.1 kaikkien parametrien osalta. Konvergenssi varmistettiin näiden potilaiden osalta myös graafisin tarkasteluin.

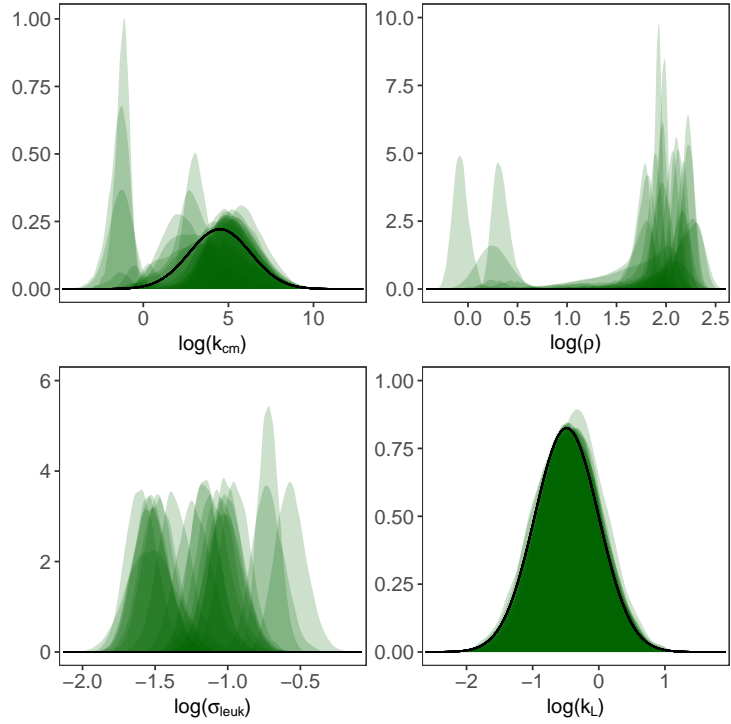
Kaikkien parametrien efektiiviset otoskoot olivat potilaittain yhdistetyissä MCMC-ketjuissa joitakin poikkeuksia lukuun ottamatta suuria (minimi: 174, Q_1 : 5044, Med : 11076, Q_3 : 12306, maksimi: 13177).

5.5 Potilaiden posteriorijakaumat

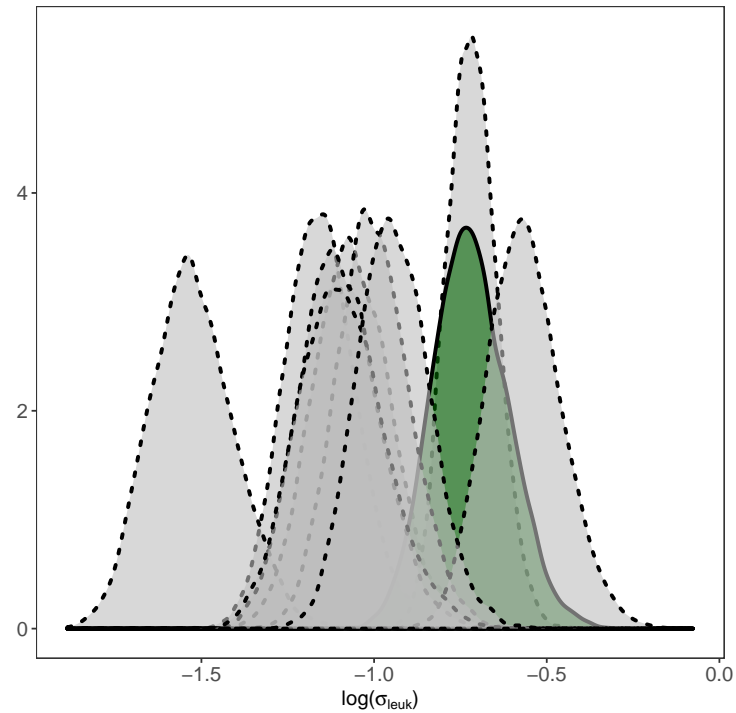
Kuva 3 esittää bayesilaisen differentiaaliyhtälömallin soviteen potilaan 15 kaksihuippuisen posteriorijakauman eri huippuja näytteistä MCMC-ketjuista poimitujen näytteiden perusteella. Näytteiden perusteella lasketut soviteen arvot eivät graafisesti tarkasteltuna eroa toisistaan kovinkaan paljoa. Posteriorijakauman kahden huipun välillä parametrien marginaalijakaumat erosivat lähinnä parametrien k_{cm} ja ρ välillä. Myös usealla



Kuva 3: Simulaatio 250 näytteellä potilaan 15 posteriorijakaumasta. Vaaka-akselilla on hoitopäivä ja pystyakselilla leukosyyttipitoisuus. Ylemmässä ja alemmassa kuvassa olevat näytteet on poimittu ketjuista, jotka konvergoivat potilaan 15 posteriorijakauman eri huippuihin. Kukaan kuvaan piirretty käyrä kuvaa mallin sovitea tietyllä posteriorijakaumasta poimitulla näytteellä.



Kuva 4: Potilaskohtaisten marginaaliposteriorijakaumien tiheysfunktioestimaatit (vihreällä) ja logaritmis-kaalassa olevat priorijakaumat (mustalla) parametreille k_{cm} , ρ , σ_{leuk} ja k_L .



Kuva 5: Parametrin σ_{leuk} marginaaliposteriorijakaumat logaritmis-kaalassa niiden potilaiden osalta, joille on tehty TPMT-genotyypin määritys. TPMT-villityypin määritys (TPMT-variant type) olevien potilaiden jakaumat on piirretty kuvaan harmaalla ja katkoviivalla. TPMT-heterotsygootin potilaan 5 jakauma on piirretty vihreällä ja yhtenäisellä viivalla.

potilaalla, jolla parametrien posteriorijakauma oli yksihiippuinen, parametrien k_{cm} ja ρ välillä havaittiin (logaritmiskaalassa) korkeahkoja korrelaatioita (minimi: 0.494, Q_1 : 0.595, Md : 0.696, Q_3 : 0.797, maksimi: 0.898).

Kuva 4 esittää logaritmiskaalassa parametrien k_{cm} , ρ , σ_{leuk} ja k_L marginaaliposteriorijakaumat potilaittain. Potilaat 9 ja 16, joilla MCMC-ketjujen konvergenssia ei saavutettu, eivät ole kuvan tarkastelussa mukana. Kuvassa näkyvät tummanvihreät alueet vastaavat niitä parametrien arvoja, joissa usean potilaan marginaaliposteriorijakaumalla on todennäköisyysmassaa. Musta käyrä on taulukon 9 mukaisen priorijakauman tiheysfunktio. Parametri L_{init} puuttuu kuvasta, sillä sen priorijakauma riippuu potilaan ensimmäisestä leukosyyttimittauksesta ja ei siten ole kaikille potilaille sama.

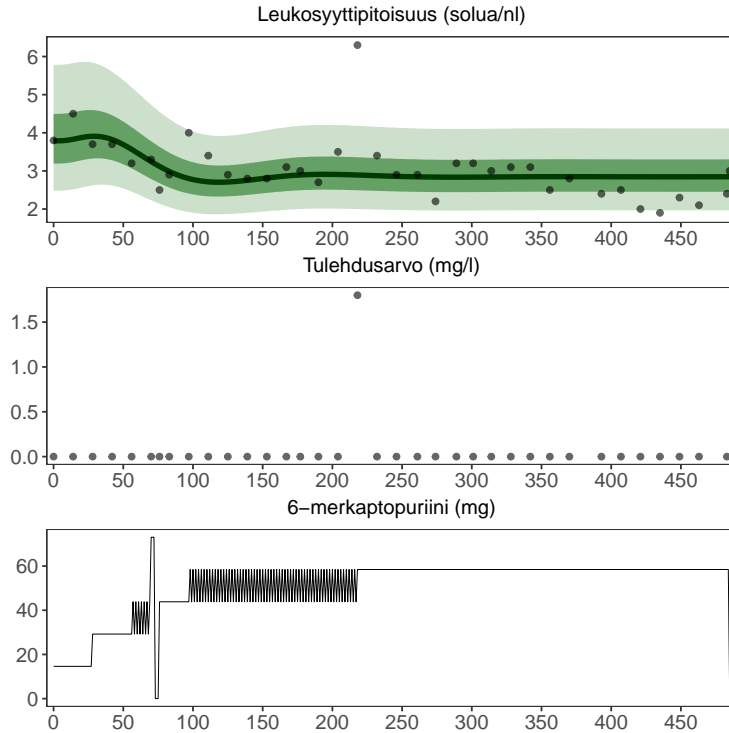
Kuvassa erottuu selkeästi erityisesti parametrien k_{cm} ja ρ marginaaliposteriorijakaumissa näkyvät kaksihiippuisuudet. Lisäksi parametrin k_L marginaaliposteriorijakauma on kullekin potilaalle käytännössä sama kuin parametrin priorijakauma, joten aineisto ei sisällä juurikaan informaatiota tästä parametrusta. Myös Friberg ym. [2002] raportoivat, että parametri k_L ei identifioitunut kovin hyvin malliversiossa, jossa parametri estimoitii. Parametrin σ_{leuk} marginaaliposteriorijakaumissa nähdään olevan vaihtelua potilaiden välillä. Tämä vaihtelu kertoo siitä, että eri potilaiden havaituista leukosyyttipitoisuuksista jää selittämättä vaihtelevia määriä. Syynä tälle voi olla esimerkiksi se, että malli ei huomioi kaikkia ilmiön kannalta oleellisia tekijöitä. Kuvassa 5 tätä vaihtelua on tarkasteltu potilaan TPMT-genotyypin suhteen niiden potilaiden osalta, joilta TPMT-genotyyppi on mitattu. Kuvan 5 (ja 4) tarkastelu paljastaa, että TPMT-heterotsygootin potilaan 5 havaintojen keskihajonta on suurehko useaan muuhun aineiston potilaaseen verrattuna.

5.6 Mallin mukaiset leukosyyttipitoisuudet

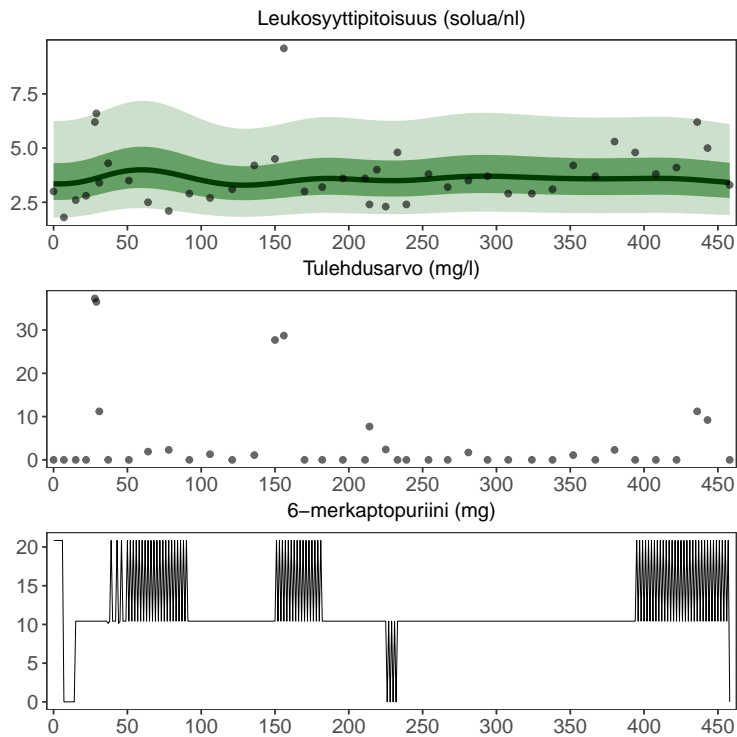
Kuvat 6 - 9 esittävät mallin mukaisen havaintojen jakauman neljän aineiston potilaan osalta leukosyyttipitoisuuden, tulehdusarvon sekä annetun 6MP:n suhteen. Kuviin on valittu hoidonaikaisten tulehdusten määrän suhteen erilaisia potilaita. Kuvissa on näkyvissä mallin mukaisen havaintojen jakauman kvartiiliväli sekä keskimäinen 90 % jakaumasta. Fraktiilit on laskettu poimimalla potilaiden posteriorijakaumista 500 näytettä ja ratkaisemalla tilayhtälö (10) kullakin näytteellä ja tiedoilla potilaan lääkekuurista. Mallin mukainen havaintojen jakauma hetkellä t on normaalijakaumien sekoitusjakauma, jonka komponenttijakaumien odotusarvot ovat näytteitä vastaavat sovittien arvot hetkellä t ja keskihajonnat vastaavat kussakin poimitussa näytteessä olevaa parametrin σ_{leuk} arvoa. Ajanhetkittäisten sekoitusjakaumien fraktiilit on laskettu numeerisesti ajanhetkittäin, muunnettu eksponenttimuunnoksella lineaariskaalaan ja piirretty kuviin. Kuvia tarkastellessa on syytä huomata, että pystyakselin skaalat vaihtelevat kuvien välillä.

Kuvan 6 perusteella mallinnus potilaan 2 aineistossa on onnistunut hyvin. Mallin mukaisen havaintojen jakauman keskimäinen 90 % sisältää yhtä havaintoa lukuun ottamatta kaikki potilaan leukosyyttimittaukset. Yksi poikkeava havainto havaitaan kuitenkin noin päivänä 225, ja se on selvästi yhteydessä samaan aikaan korkealla olleeseen tulehdusarvoon. Tätä yksittäistä korkeaa tulehdusarvomittausta lukuun ottamatta potilaan elimistössä ei ole ollut tulehduksia hoidon aikana. Kuvassa kiinnostavaa on myös hoidon loppupuolella havaittava mallin yliennustaminen, joka ei selity lääkityksen muutoksella, sillä potilaan lääkitys on ollut vakio noin päivästä 225 lähtien.

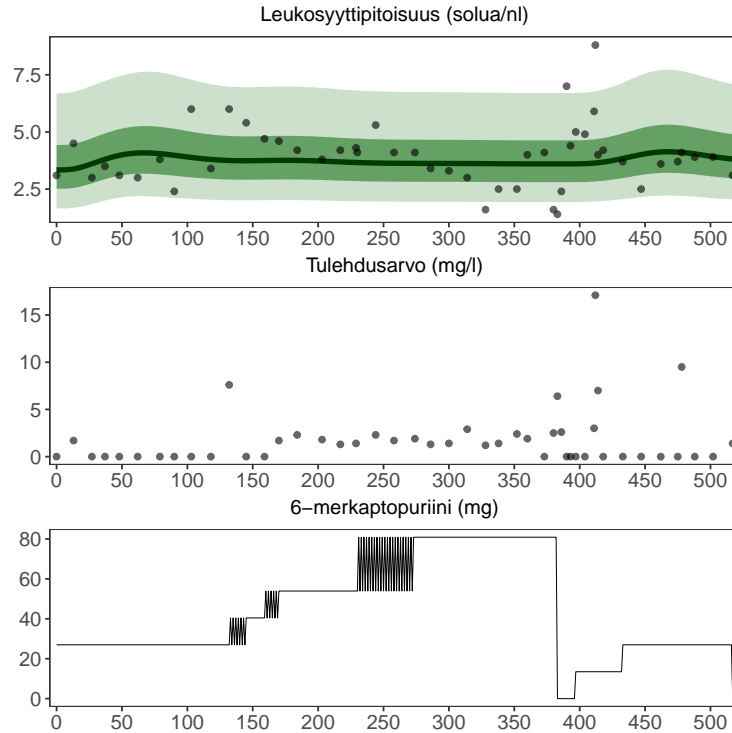
Kuva 7 esittelee mallinnustuloksia potilaan 3 aineistossa. Potilaaseen 2 verrattuna potilaalla 3 on ollut hieman enemmän tulehduksia hoidon aikana. Tulehdustilat eivät kui-



Kuva 6: Bayesilaisen differentiaaliyhtälömallin sovite potilaan 2 aineistoon. Vaaka-akselilla on hoitopäivä. Mustat pisteet ovat havaintoja. Ylimpänä havaitut leukosyyttipitoisuudet, keskellä havaittu tulehdusarvo ja alimpana annettu lääkitys. Sahalaitakuvio lääkityksen kuvassa osoittaa, että annosta on vaihdeltu vuoropäivin. Leukosyyttipitoisuuden aikasarjaan on piirretty ajanhetkittäin tummanvihreällä kvartiiliväli ja vaalean vihreällä keskimäinen 90 % mallin mukaisesta leukosyyttipitoisuuden jakaumasta. Musta viiva kuvaa mallin mukaisen havaintojen jakauman mediaania.



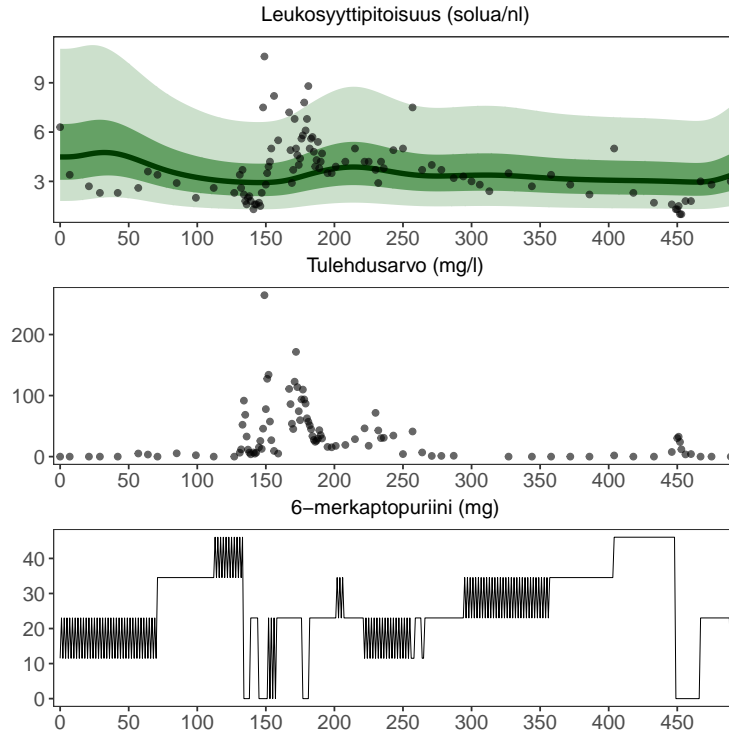
Kuva 7: Bayesilaisen differentiaaliyhtälömallin sovite potilaan 3 aineistoon. Kuvateksti kuten kuvassa 6.



Kuva 8: Bayesilaisen differentiaaliyhtälömallin sovite potilaan 6 aineistoon. Kuvateksti kuten kuvassa 6.

tenkaan ole jatkuneet kovin pitkiä aikoja, vaan näyttävät kuvan perusteella maksimissaan muutaman päivän mittaisilta. Liki pitäen päivien 25, 75, 150, 225, 375 ja 425 kohdalla havaitaan mallin mukaisen kvartiilivälin ulkopuolella olevia leukosyyttihavaintoja, joilla on kuvan perusteella yhteys samoihin aikoihin koholla olleisiin tulehdusarvoihin. Mallinnus näyttää onnistuneen potilaan aineistossa kohtalaisesti.

Mallinnus ei ole onnistunut potilaan 6 aineistossa yhtä hyvin kuin potilaiden 2 ja 3 aineistoissa (kuva 8). Lukuun ottamatta noin päivien 75 – 150 välistä aikaa, mallin mukainen havaintojen jakauma näyttää sopivan havaittuihin leukosyyttipitoisuuksiin kohtalaisesti noin hoitopäivään 300 asti. Päivien 75 – 150 välillä esiintyvät poikkeamat selittynevät osittain noin päivänä 125 mitatulla yksittäisellä korkeahkolla tulehdusarvomittauksella. Noin päivään 300 asti kasvatettu lääkeannos ei kuitenkaan näy mallin mukaisessa havaintojen jakaumassa juuri lainkaan. Kuvan perusteella vaikuttaa siltä, että noin päivän 300 kohdalla 80 milligrammaan nostetun lääkeannoksen tulisi vaikuttaa mallin mukaiseen leukosyyttipitoisuuteen voimakkaammin, sillä tällöin potilaan havaitut leukosyyttipitoisuudet näyttävät putoavan aiempaa matalammalle tasolle. Tämän jälkeen potilaan lääkitys on hetkellisesti lopetettu kokonaan, jolla ei ole kuvasta katsottuna myöskään vaikutusta mallin mukaiseen havaintojen jakaumaan. Myös potilaan tulehdusarvo on ollut koholla jo noin päivästä 175 alkaen ja saavuttaa huippunsa suurin piirtein samaan aikaan lääkityksen keskeyttämisen kanssa. Mallin yli- ja aliennustaminen päivän 300 jälkeen on todennäköisesti seurausta siitä, että mallin mukainen havaintojen jakauma asettuu keskiarvoistaen havaitun matalan ja korkean leukosyyttipitoisuuden keskelle, sillä malli ei ole löytänyt selkeää yhteyttä lääkityksen ja leukosyyttipitoisuuden välille. Mikäli malli reagoisi paremmin potilaan lääkityksen muutoksiin ja potilaan tulehdusarvo olisi huomioitu mallissa, mallin mukainen havaintojen jakauma saattaisi näyttää erilaiselta erityisesti hoitopäivien 350 – 450 välillä.



Kuva 9: Bayesilaisen differentiaaliyhtälömallin sovite potilaan 4 aineistoon. Kuvateksti kuten kuvassa 6.

Malli toimii heikosti myös potilaan 4 aineistossa (kuva 9). Aineisto on haastava, sillä potilaalla on ollut pitkäkestoinen tulehdus aikavälillä 125 – 250 päivää. Aikavälillä mitatut tulehdusarvot ovat selvästi yhteydessä samaan aikaan havaittuihin leukosyyttipitoisuuksiin. Malli on selittänyt tulehduksesta johtuvan leukosyyttipitoisuuden vaihtelun havaintojen keskihajonnalla, koska hoidon aikaisia tulehduksia ei ole huomioitu mallissa. Tämän vuoksi mallin mukainen havaintojen jakauma on leveä myös silloin, kun potilaan elimistössä ei ole ollut tulehdusta. Noin päivien 150 – 175 aikana mallin mukaisen havaintojen jakauman 90 %:n fraktiili ei vaikuta realistiselta, sillä sen yläpuolella nähdään olevan useita havaintoja. Siten vaikuttaa siltä, että mallin arviot leukosyyttipitoisuuden epävarmuudesta eivät ole hyviä tulehduksen aikana. Hoidon alussa ja lopussa tapahtuva mallin yliennustaminen johtunee todennäköisesti siitä, että malli on pääasiassa sovittunut noin päivien 125 – 300 aikaisiin mittauksiin. Potilaan lääkityksen huomataan lisäksi muuttuneen hoidon aikana useita kertoja, mutta tämä ei kuvan 8 tapaan juuri näy mallin mukaisessa havaintojen jakaumassa.

Kuvien 6 - 9 perusteella vaikuttaa siltä, että bayesilainen differentiaaliyhtälömalli toimii kohtalaisesti hoitohistorioissa, joissa potilaalla on ollut enintään lyhytaikaisia tulehdusjaksoja. Jos potilaalla on ollut pitkäaikaisia tulehduksia, malli ei toimi hyvin. Lisäksi mallin mukaiset leukosyyttipitoisuudet vaikuttavat reagoivan heikosti 6MP-annostuksen muutoksiin. Mallin pohdintaan palataan kappaleessa 7.1.

6 Bayesilainen epälineaarinen tila-avaruusmalli

Tässä kappaleessa esitellään työn toinen malli ja siihen liittyvät mallinnustulokset. Myös mallin parametrien estimointia sekä estimoinnin diagnostiikkaa tarkastellaan. Esiteltävä malli on epälineaarinen tila-avaruusmalli, joka kehittää kappaleen 5 mallia lisäämällä tilayhtälön leukopoeesimalliin satunnaisuutta, jota mallinnetaan potilaan tulehdusarvon perusteella. Malli yksinkertaistaa myös 6MP:n farmakokinetiikan mallintamista farmakokineettiseen malliin (7) verrattuna.

Epälineaarilla tila-avaruusmallilla tarkoitetaan mallia, joka muodostuu kahdesta yhtälöstä (Durbin ja Koopman [2012], s. 224 – 227):

$$\begin{aligned}\alpha_t &= T_t^{(\theta)}(\alpha_{t-1}) + R_t^{(\theta)}(\alpha_{t-1})\eta_t \quad (\text{tilayhtälö}) \\ \alpha_0 &\sim N(\mu_0^{(\theta)}, \Sigma_0^{(\theta)})\end{aligned}\tag{13}$$

$$y_t = Z_t^{(\theta)}(\alpha_t) + H_t^{(\theta)}(\alpha_t)\epsilon_t. \quad (\text{havaintoyhtälö})$$

Edellä virheet $\eta_t \sim N(0, I)$ ja $\epsilon_t \sim N(0, I)$ ovat riippumattomia ajallisesti aiemmista virheistä ja toisistaan. Funktiot $T_t^{(\theta)}$, $R_t^{(\theta)}$, $Z_t^{(\theta)}$ ja $H_t^{(\theta)}$ ovat moniulotteisia sekä mahdollisesti ajasta riippuvia, ja ne oletetaan differentioituviksi. Tilayhtälön alkutila α_0 noudattaa moniulotteista normaalijakaumaa odotusarvolla $\mu_0^{(\theta)}$ ja kovarianssimatriisilla $\Sigma_0^{(\theta)}$. Merkin­­töjen yläindeksointi parametrivektorilla θ viittaa mahdolliseen riippuvuuteen mallin parametreista. Epälineaarisen tila-avaruusmallin teoria voidaan yleistää myös tilanteeseen, jossa tila- ja havaintoyhtälössä esiintyvät virheet eivät ole normaalijakautuneet, mutta tässä työssä rajoitutaan normaalijakautuneiden virheiden tapaukseen.

Havaintoyhtälössä (*observation equation*) esiintyvä vektori y_t on p -ulotteinen ja sisältää hetkellä $t \in \{1, \dots, n\}$ havaittavan havainnon (Durbin ja Koopman [2012], s. 43 – 44). Vektoria y_t kutsutaan mallin havaintovektoriksi (*observation vector*). Havaintovektorilla viitataan tässä työssä sekä havaintoon hetkellä t että mallin havaintovektoriin ylipäätään kiinnittämättä ajanhetkeä t . Sama pätee tilayhtälössä (*state equation*) esiintyvälle m -ulotteiselle vektorille α_t , jota kutsutaan mallin tilavektoriksi (*state vector*).

Tila- ja havaintoyhtälöt kuvaavat, kuinka tilavektori ja havaintovektori muuttuvat ajassa ja riippuvat toisistaan. Epälineaarissa tila-avaruusmallissa tutkittavan ilmiön tilaa kuvaavaa tilavektoria α_t ei havaita, vaan tiedossa ovat ainoastaan havainnot y_t hetkillä $t = 1, \dots, n$, joiden oletetaan riippuvan tilavektorista havaintoyhtälön mukaisesti. Tavoitteena mallinnuksessa on tehdä tuntematonta tilavektoria koskevaa tilastollista päättelyä havaintojen perusteella. Myös parametrivektori θ täytyy estimoida. Estimointia käsitellään kappaleessa 6.3.

6.1 Mallin kuvaus

Bayesilaisella epälineaarilla tila-avaruusmallilla viitataan jatkossa tässä kappaleessa esitettävään mallikokonaisuuteen. Tila- ja havaintoyhtälöstä muodostuva malli esitetään aluksi kokonaisuutena, jonka jälkeen mallin yksityiskohdat käsitellään kappaleissa 6.1.1–6.1.4. Malli sovitetaan kunkin potilaan aineistoon erikseen, joten merkintöjen yksinkertaistamiseksi kyseessä olevaa potilasta ei enää jatkossa indeksoida järjestyksellä i . Mallin tila- ja havaintoyhtälö on esitetty kaavoissa (14) ja (15), tässä järjestyksessä.

$$\begin{aligned}
M_k &= \left(M_{k-1} - e_{tgn} \frac{D_k}{B_k} \right) \exp(-k_{tgn} \Delta t_k) + e_{tgn} \frac{D_k}{B_k} \\
V_k &= V_{k-1} \exp(-\theta_{ou} \Delta t_k) + \frac{\sigma_{ou}}{\sqrt{2\theta_{ou}}} \exp(-\theta_{ou} \Delta t_k) \left(\sqrt{\exp(2\theta_{ou} \Delta t_k) - 1} \right) \eta_{1,k} \\
S_k &= S_{k-1} + \left[\left(1 - \frac{E_{max} M_{k-1}}{E_{C50} + M_{k-1}} \right) \left(\frac{L_0}{L_{k-1}} \right)^\gamma - 1 \right] k_{tr} \Delta t_k + \sigma_{wn}(V_{k-1}) \sqrt{\Delta t_k} \eta_{2,k} \\
C_k^{(1)} &= C_{k-1}^{(1)} + \left[\exp(S_{k-1} - C_{k-1}^{(1)}) - 1 \right] k_{tr} \Delta t_k + \sigma_{wn}(V_{k-1}) \sqrt{\Delta t_k} \eta_{3,k} \\
C_k^{(2)} &= C_{k-1}^{(2)} + \left[\exp(C_{k-1}^{(1)} - C_{k-1}^{(2)}) - 1 \right] k_{tr} \Delta t_k + \sigma_{wn}(V_{k-1}) \sqrt{\Delta t_k} \eta_{4,k} \\
C_k^{(3)} &= C_{k-1}^{(3)} + \left[\exp(C_{k-1}^{(2)} - C_{k-1}^{(3)}) - 1 \right] k_{tr} \Delta t_k + \sigma_{wn}(V_{k-1}) \sqrt{\Delta t_k} \eta_{5,k} \\
L_k &= L_{k-1} + \left[\exp(C_{k-1}^{(3)} - L_{k-1}) - 1 \right] k_{tr} \Delta t_k + \sigma_{wn}(V_{k-1}) \sqrt{\Delta t_k} \eta_{6,k}
\end{aligned} \tag{14}$$

$$\begin{aligned}
l_k &= L_k + \epsilon_l \\
v_k &= V_k + \epsilon_v
\end{aligned} \tag{15}$$

Tila- ja havaintoyhtälö on kirjoitettu epälineaarisen tila-avaruusmallin (13) matriisi- muodosta poiketen alkioittain. Perusrakenteeltaan kuvan 2 mukainen tilayhtälö (14) muodostuu yksinkertaistetusta 6MP:n farmakokineettisestä mallista (tilamuuttujan M päivityskaava), mallista potilaan elimistössä olevalle tulehdukselle (tilamuuttujan V päivityskaava) sekä leukopoeiesimallin (1) johdannaisesta (tilamuuttujien $S - L$ päivityskaavat). Mallin havaintoyhtälö (15) liittyy potilaan havaitun leukosyyttipitoisuuden ja tulehdusarvon tilayhtälön tilamuuttujiin.

Tilayhtälössä esiintyvät virheet $\eta_{i,k}$ noudattavat jakaumaa $N(0, 1)$ kaikilla $i = 1, \dots, 6$. Symboli Δt_k kuvaa kulunutta aikaa hetkien k ja $k - 1$ välillä. Muut merkinnät ja parametrit ovat kuten aiemmin tai selitetään myöhemmin tässä kappaleessa.

6.1.1 Yksinkertaistettu 6MP:n farmakokineettinen malli

TGN-konsentraatiota kuvaavan tilamuuttujan M päivityskaava on saatu ratkaisemalla $M_k = M(t_k)$ analyttisesti differentiaaliyhtälöstä

$$\begin{aligned}
\frac{dM(t)}{dt} &= \phi(D_k, B_k) - k_{tgn} M(t) \\
\phi(D_k, B_k) &= e_{tgn} k_{tgn} \frac{D_k}{B_k} \\
M(t_{k-1}) &= M_{k-1} \\
t &\in [t_{k-1}, t_k].
\end{aligned} \tag{16}$$

Edellä D_k ja B_k ovat potilaalle annettu 6MP-annos ja potilaan kehon pinta-ala välillä $[t_{k-1}, t_k]$. Parametrit e_{tgn} ja k_{tgn} skaalaavat potilaan kehon pinta-alaan suhteutettua lääkitystä ja vaikuttavat TGN:n poistumisnopeuteen, tässä järjestyksessä.

Differentiaaliyhtälö (16) ei ota kantaa siihen, mistä TGN-konsentraation mittaaminen on tehty. Yhtälön tulkinta on, että TGN-konsentraation $M(t)$ muutosnopeuteen vaikuttaa

hetkellä $t \in [t_{k-1}, t_k)$ elimistössä oleva TGN-konsentraatio sekä funktiolla ϕ muunnettu potilaalle annettu lääkeannos. Funktion ϕ arvo kuvaa TGN-konsentraation muutosnopeuden kasvua, joka aiheutuu potilaalle annetusta lääkityksestä. Annettu lääkitys suhteutetaan funktiossa kehon pinta-alaan, koska sillä on yhteys veren tilavuuteen (Raes ym. [2006]). Veren tilavuus vaikuttaa 6MP:n imeytyttyä lääkkeen konsentraatioon veressä.

Lääkeannos ja kehon pinta-ala oletetaan (jatkuvassa ajassa) porrasfunktioiksi siten, että lääkeannos on vakio vuorokauden ajan lääkkeen antamisesta ja kehon pinta-ala on vakio kullakin tarkasteluvälillä $[t_{k-1}, t_k)$. TGN-konsentraation oletetaan olevan hoidon aluksi 0, ts. $M(0) = 0$. Näin ollen tilamuuttujan M alkujakauma on $N(0, 0)$. Bayesilainen epälineaarinen tila-avaruusmalli olettaa siten, että hoidon alussa potilaan elimistössä ei ole leukoioeesiin vaikuttavaa TGN:ää.

Differentiaaliyhtälöstä (16) ratkaistun farmakokineettisen mallin hyödyntämiseksi potilaan kehon pinta-alan on oltava tunnettu. Kehon pinta-alaa neliömetreissä on mahdollista arvioida käyttämällä nk. Mostellerin kaavaa (Mosteller [1987], El Edelbi ym. [2011]):

$$\sqrt{\frac{hw}{3600}}, \quad (17)$$

jossa h on pituus senttimetreinä ja w paino kilogrammoissa. Kuten kappaleessa 3 todettiin, potilaiden painoista ja pituuksista on aineistossa tieto kuitenkin vain hoidon alkumittauksesta. Tämän vuoksi potilaiden hoidon aikaista pituuskasvua ja painonnousua arvioitiin lasten kasvutilastojen perusteella [Centers for Disease Control and Prevention, 2017]. Käytetyt tilastot sisälsivät yhdeksän fraktiilin avulla ilmoitetun jakauman alle 20-vuotiaiden tyttöjen ja poikien painoista ja pituuksista kuukauden tarkkuudella. Potilaiden sukupuolen ollessa kappaleeseen 3 viitaten tuntematon, poikien ja tyttöjen paino- ja pituustilastot yhdistettiin keskiarvoistamalla vastaavien kuukausien fraktiilit kummassakin tilastossa. Tämän jälkeen jokaisen potilaan pituuskasvua ja painonnousua arvioitiin hyödyntämällä lineaarista interpolointia ja oletusta, että potilaan pituus ja paino kasvavat tilastojen mediaanikasvun mukaisesti. Interpolointi aloitettiin iästä, jonka mukainen mediaanipituus ja -paino osui lähimmäksi potilaan alkumittauksessa mitattua pituutta ja painoa. Likimääräinen potilaskohtainen kehon pinta-ala hoidon aikana saatiin edellä kuvatun menettelyn jälkeen laskettua kaavan (17) avulla.

6.1.2 Malli tulehdusarvolle

Tilamuuttuja V kuvaa tilayhtälössä (14) potilaan elimistössä olevaa tulehdusta. Potilaan havaittu tulehdusarvo on mallinnettava tilamuuttujan V avulla, sillä käytössä olevien hoitohistorioiden tulehdusarvoaikasarjoissa on puuttuvaa tietoa. Bayesilainen epälineaarinen tila-avaruusmalli olettaa tilamuuttujan V noudattavan odotusarvoon hakeutuvaa Ornstein-Uhlenbeck-prosessia¹¹ (*mean reverting Ornstein-Uhlenbeck process*) (Brigo ja Mercurio [2006], s. 58 – 59). Prosessin määrittelee stokastinen differentiaaliyhtälö:

$$\begin{aligned} dV_t &= \theta_{ou}(\mu_{ou} - V_t)dt + \sigma_{ou}dB_t^{(V)} \\ V_0 &= v_0, \end{aligned} \quad (18)$$

¹¹ Finanssialalla tätä mallia kutsutaan Vasicek-malliksi ja sitä käytetään usein korkotasojen mallintamisessa.

jossa v_0 on prosessin alkuarvo, θ_{ou} sekä σ_{ou} ovat parametreja ja $B_t^{(V)}$ on standardi Brownin liike. Prosessia noudattavan satunnaismuuttujan V_k ehdollinen jakauma ehdolla edellinen havainto on normaalijakauma, jonka odotusarvo ja varianssi ovat (Franco [2003], liitemateriaali):

$$\begin{aligned} E[V_k|v_{k-1}] &= \mu_{ou} + (v_{k-1} - \mu_{ou}) \exp(-\theta_{ou}\Delta t_k) \\ Var[V_k|v_{k-1}] &= \frac{\sigma_{ou}^2}{2\theta_{ou}} [1 - \exp(-2\theta_{ou}\Delta t_k)]. \end{aligned} \quad (19)$$

Tilayhtälössä (14) oleva tilamuuttujan V päivityskaava seuraa kaavan (19) perusteella, kun $\mu_{ou} = 0$.

Perustelu odotusarvoon hakeutuvan Ornstein-Uhlenbeck-prosessin käytölle tulehduksen mallintamisessa on se, että tulehdusarvo näyttää graafisesti tarkasteltuna (kuva 1) hakeutuvan nolnaan tulehdustilojen välillä. Huomattavaa on, että tilamuuttujan V mallintamisen tarkoituksena ei ole ennustaa, milloin potilaan elimistössä alkaa tulehdustila, vaan kuvata tulehduksen kehitystä ajassa tulehdusarvomittausten välillä. Parametrin θ_{ou} suuruus vaikuttaa käänteisesti nopeuteen, jolla prosessi hakeutuu kohti odotusarvoaan, ja parametri σ_{ou} vaikuttaa prosessin odotusarvon ympärillä tapahtuvan vaihtelun suuruuteen [Franco, 2003]. Tilamuuttujan V alkujakaumaksi oletetaan $N(v_1, 0)$, jossa v_1 on potilaan ensimmäinen (muunnettu) tulehdusarvomittaus. Tulehdusarvomittausten muunnoksesta kerrotaan tarkemmin kappaleessa 6.1.4. Ensimmäisen tulehdusarvomittauksen puuttuessa tilamuuttujan V alkujakauman oletetaan olevan $N(0, 0)$.

6.1.3 Tulehduksen liittäminen leukopoeesimalliin

Leukopoeesia kuvaavien tilamuuttujien $S - L$ päivityskaavat tilayhtälössä (14) perustuvat leukopoeesimallin (1) (logaritmiskaalassa) muuntamiseen stokastiseksi differentiaaliyhtälöryhmäksi (20), jossa jokaisen tilamuuttujan differentiaaliyhtälöön on lisätty valkoista kohinaa stokastisen differentiaaliyhtälön (18) tilamuuttujasta V_t riippuvalla intensiteetillä $\sigma_{wn}(V_t)$.

$$\begin{aligned} dS_t &= \left(k_{tr} \left(1 - \frac{E_{max}M_t}{E_{C50} + M_t} \right) \left(\frac{L_0}{L_t} \right)^\gamma - k_{tr} \right) dt + \sigma_{wn}(V_t)dB_t^{(S)} \\ dC_t^{(1)} &= \left(k_{tr} \exp(S_t - C_t^{(1)}) - k_{tr} \right) dt + \sigma_{wn}(V_t)dB_t^{(C^{(1)})} \\ dC_t^{(2)} &= \left(k_{tr} \exp(C_t^{(1)} - C_t^{(2)}) - k_{tr} \right) dt + \sigma_{wn}(V_t)dB_t^{(C^{(2)})} \\ dC_t^{(3)} &= \left(k_{tr} \exp(C_t^{(2)} - C_t^{(3)}) - k_{tr} \right) dt + \sigma_{wn}(V_t)dB_t^{(C^{(3)})} \\ dL_t &= \left(k_{tr} \exp(C_t^{(3)} - L_t) - k_{tr} \right) dt + \sigma_{wn}(V_t)dB_t^{(L)}. \end{aligned} \quad (20)$$

Stokastisessa differentiaaliyhtälöryhmässä (20) $B_t^{(S)}, \dots, B_t^{(L)}$ ovat toisistaan (sekä Brownin liikkeestä $B_t^{(V)}$) riippumattomia standardeja Brownin liikkeitä ja σ_{wn} on tilamuuttujan V_t funktio. Funktio σ_{wn} on valittu siten, että

$$\sigma_{wn}(V_t) = \exp(\beta_{lp}) \exp(\beta_{vol}V_t) = \sigma_{lp} \exp(\beta_{vol}V_t). \quad (21)$$

Tämä tekee bayesilaisesta epälineaarista tila-avaruusmallista volatilitteettimallin, jossa leukopoeesiin liittyvä epävarmuus riippuu siitä, kuinka suuri tulehdus potilaan elimistössä

on (Durbin ja Koopman [2012], s. 216 – 217). Kaavassa (21) parametri σ_{lp} kuvaa leukopieesiin liittyvien tilamuuttujien virheiden keskihajontaa tilanteessa, jossa tilamuuttujan V_t arvo on 0. Parametri β_{vol} säätelee, kuinka paljon yhden yksikön muutos tilamuuttujassa V_t vaikuttaa tilamuuttujien $S - L$ keskihajonnan logaritmiin.

Tilamuuttujien $S - L$ päivityskaavat on muodostettu tilayhtälössä (14) ratkaisemalla stokastinen differentiaaliyhtälöryhmä (20) approksimatiivisesti Euler-Maruyama menetelmällä (Kloeden ja Platen [1995], s. 305 – 307). Diskreetointiväliksi Δt_k asetettiin Euler-Maruyama menetelmässä 0.1 päivää, jota varten aineiston potilaiden hoitohistorioiden mittausväli muunnettiin 0.1 päivään. Tilamuuttujien $S - L$ alkujakaumiksi oletetaan $N(l_1, 1)$. Tällä oletuksella stokastinen differentiaaliyhtälöryhmä (20) on tilamuuttujaa S_t lukuun ottamatta keskimäärin tasapainotilassa hoidon alkaessa.

6.1.4 Havaintoyhtälö

Havaintoyhtälössä (15) l_k on logaritmoitu leukosyyttipitoisuus, kuten aiemmin, ja v_k on muunnoksella $\log(x + 1)$ muunnettu tulehdusarvohavainto. Tulehdusarvohavaintojen muunnokseksi valittiin $\log(x + 1)$, sillä tämä muunnos säilyttää arvoltaan nolla olevat tulehdusarvohavainnot nolliina. Lisäksi muunnoksen käyttäminen skaalaa poikkeavimpia tulehdusarvoja käyttökelpoisemmalle skaalalle. Havaintovirheille ϵ_l ja ϵ_v oletetaan, että $\epsilon_l \sim N(0, \sigma_{leuk}^2)$ ja $\epsilon_v \sim N(0, \sigma_{crp}^2)$.

6.2 Parametrien priorijakaumat ja kiinnitetyt arvot

Bayesilaisen epälineaarisen tila-avaruusmallin parametrien priorijakaumat ja kiinnitetyt arvot on koottu taulukkoon 10. Priorijakaumien valinnalla pyrittiin pääasiassa rajaamaan pois mahdollisia parametrien arvoja.

Parametrin γ arvo asetettiin taulukossa 3 olevaan keskimääräiseen arvoon. Parametri E_{max} asetettiin arvoon 1, joka on luonnollinen valinta tilamuuttujan S päivityskaavassa olevan termin

$$1 - \left(\frac{E_{max}M}{E_{C50} + M} \right)$$

perusteella. Kun $E_{max} = 1$, termi saa arvon 0 tilamuuttujan M lähestyessä ääretöntä, ja arvon 1, kun tilamuuttuja $M = 0$.

Havaintoyhtälön (15) hajontaparametrit σ_{leuk} ja σ_{crp} kiinnitettiin arvoihin $\sigma_{leuk} = 0.057$ ja $\sigma_{crp} = 0.1$. Amundsen ym. [2012] mukaan erään neutrofiilien laskennassa käytettävän mittalaitteen mittaustarkkuuden variaatiokerroin on noin 0.057. Arvoa voidaan käyttää myös leukosyytien tapauksessa, sillä artikkelin mukaan solupitoisuuksien mittaustarkkuus riippuu pääosin pitoisuuden suuruudesta, jota mitataan. Näin ollen tiettyä neutrofiilipitoisuutta vastaava leukosyyttipitoisuuden mittaustarkkuus on vähintäänkin sama tai suurempi, koska neutrofiilit ovat osajoukko leukosyyteistä. Luku 0.057 on kahdesta edellä mainitussa artikkelissa käsitellystä laitteesta epätarkemman keskimääräinen mittaustarkkuuden variaatiokerroin, kun neutrofiilipitoisuus on välillä 0 – 2 solua/nl. Koska mallin tilamuuttuja L on logaritmiskaalassa, variaatiokerrointa voidaan käyttää suoraan parametrin σ_{leuk} arvona. Tämä voidaan nähdä tarkastelemalla lineaariskaalassa olevaa normaalijakautunutta leukosyyttimittausta y_k , jonka odotusarvo ja varianssi ovat Y_k ja σ^2 ja joka muunnetaan logaritmiskaalaan:

$$\begin{aligned}
y_k &= Y_k + \epsilon_k \\
\iff y_k &= Y_k + Y_k \eta_k \\
\implies l_k &= \log(Y_k) + \log(1 + \eta_k) \\
\implies l_k &\approx \log(Y_k) + \eta_k.
\end{aligned}$$

Edellä siis $\epsilon_k \sim N(0, \sigma^2)$ ja siten $\eta_k \sim N\left(0, \frac{\sigma^2}{Y_k^2}\right)$. Viimeisellä rivillä käytetty approksimaatio perustuu siihen, että $\log(1 + x) \approx x$, kun x on pieni. Havaintoyhtälön (15) leukosyyttipitoisuutta koskevan yhtälön havaitaan olevan viimeisen rivin mukainen, sillä $L_k = \log(Y_k)$.

Roberts ym. [2000] mukaan tulehdusarvon mittauksessa käytettävien mittalaitteiden variaatiokerroin on noin 0.1. Tilamuuttuja V ei ole logaritmisskaalassa, mutta funktio $\log(x + 1)$ on lähellä tavallista logaritmfunktiota, kun $x \geq 5$, joka on tyypillinen arvoalue tulehdusarvolle tulehduksen aikana. Täten arvon $\sigma_{crp} = 0.1$ käyttäminen on myöskin perusteltua.

Parametrin L_0 priorijakauma asetettiin käyttämällä lähtökohtana kappaleessa 2.2 mainittua 2 – 6-vuotiaiden lasten veren leukosyyttipitoisuuden vaihteluväliä 5 – 13 solua/nl. Tälle välille sopiva normaalijakauma saatiin asettamalla $\log(5)$ ja $\log(13)$ logaritmisskaalassa olevan normaalijakauman 2.5 %:n ja 97.5 %:n fraktiilien arvoiksi ja selvittämällä jakauma näiden tietojen perusteella.

Parametrin k_{tr} arvo määrätään parametrin t_{MTT} arvon perusteella, kuten kaavassa (2). Friberg ym. [2002] estimoivat taulukon 3 mukaisesti parametrille t_{MTT} arvoja väliltä 3.73 – 5.62 päivää (noin 90 – 135 tuntia). Flidner ym. [1964] mukaan t_{MTT} voi olla tulehdistilanteessa kuitenkin jopa 48 tuntia. Sen sijaan Dancey ym. [1976] mukaan t_{MTT} on noin 158 tuntia. Parametrille t_{MTT} asetettiin näiden tietojen valossa priorijakauma, jossa valtaosa todennäköisyydestä (n. 90 %) on keskittynyt arvon 6.6 päivää (158 tuntia) alapuolelle. Priorijakauman mukaan parametri t_{MTT} on 95 %:n todennäköisyydellä alle 8.5 ja 99 %:n todennäköisyydellä alle 12 päivää.

Hajontaan liittyvien parametrien β_{vol} , σ_{ou} , θ_{ou} ja σ_{lp} priorijakaumat asetettiin logaritmisskaalassa epäinformatiivisiksi. Priorijakaumien keskihajonnat pidettiin kuitenkin maltillisina, sillä näiden parametrien arvot eivät voi olla logaritmisskaalassa kovin suuria, jotta malli on järkevä.

Parametrin E_{C50} priorijakauma asetettiin tarkastelemalla Derijks ym. [2004], Hindorf ym. [2006] artikkeleissa ja Franssonin [2007] väitöskirjassa esillä olleita TGN_{RBC}-konsentraatioita. Näissä lähteissä havaitut TGN_{RBC}-konsentraatiot olivat tyypillisesti alle 1000 pikomoolia kohti $8 \cdot 10^8$ punasolua. Asetetun priorijakauman mukaan TGN hidastaa kantasolujen tuotantoa puolet maksimista konsentraatiolla, joka on 70 %:n todennäköisyydellä alle 2660 ja 90 %:n todennäköisyydellä alle 13750 pikomoolia $8 \cdot 10^8$ punasolua kohti. Parametrille E_{C50} sallittiin laaja arvoalue, sillä 6MP:n farmakokinetiikassa on suuria yksilöllisiä eroja.

Tilamuuttujan M päivityskaavassa esiintyvillä parametreille e_{tgn} ja k_{tgn} muodostettiin priorijakauma sovittamalla tilayhtälön (14) farmakokineettinen malli kappaleessa 4.2.2 mainitun Derijks ym. [2004] artikkelin kuvasta hankittuihin potilaskohtaisiin TGN_{RBC}-konsentraatioihin.

Aineisto kerättiin kuvasta käyttämällä tietokoneohjelmaa, jolla havaintopisteiden arvot saatiin poimittua riittävällä tarkkuudella. Tietojen kerääminen onnistui 26 potilaal-

ta 30:sta. Artikkelin tietojen mukaan potilaille oli annettu päivittäistä 50 milligramman 6MP-annosta ja potilaiden keskimääräinen painoon suhteutettu 6MP-annos oli 0.71 milligrammaa painokiloa kohti (95 %:n luottamusväli: [0.66, 0.76]). Lisäksi tiedossa oli, että potilaista 20 oli miehiä ja 10 naisia.

Farmakokineettisen mallin sovittamiseksi tarvittava potilaiden kehon pinta-ala saatiin arvioimalla kerätyn aineiston potilaiden keskimääräinen paino ja pituus sekä hyödyntämällä kaavaa (17). Arvio keskimääräiselle painolle saatiin olettamalla kunkin potilaan paino keskimääräiseksi ja ratkaisemalla tuntematon keskimääräinen paino edellä mainittujen lääkeannosta koskevien tietojen perusteella. Keskimääräinen pituus arvioitiin painotettuna keskiarvona hollantilaisten miesten ja naisten keskimääräisistä pituuksista¹², sillä potilaat olivat kerätyssä aineistossa hollantilaisia. Näitä tietoja hyödyntämällä potilaiden keskimääräisen kehon pinta-alan arvioksi saatiin 1.865 m^2 . Arvoa käytettiin kunkin potilaan lääkeannoksen suhteuttamiseen mallin sovittamisessa.

Mallin sovittaminen tehtiin ratkaisemalla optimointiongelma, jossa minimoitiin havaintojen ja mallin sovituksen neliöityjen etäisyyksien summaa. Mallinnuksen ulkopuolelle jätettiin yksi potilas, jonka sovitteeksi estimoitui kasvava suora sen vuoksi, että potilaan mittaukset olivat jääneet tekemättä tutkimuksen tarkastelujakson loppupuolella. Muilla potilailla mallin sovite näytti asettuvan tasapainokonsentraatioon, joka vaikuttaa järkevältä esimerkiksi Franssonin [2007] väitöskirjan sivulla 32 olevan TGN_{RBC} -konsentraatioita esittävän kuvan perusteella. Kuva 10 esittelee mallin sovitteet kolmen potilaan osalta.

Potilaskohtaisten parametrien estimoinnin jälkeen parametriestimaatteihin sovitettiin kaksiulotteinen t-jakauma kolmella vapausasteella. Jakauman sovittaminen tehtiin käyttämällä moniulotteiselle t-jakaumalle soveltuvaa EM-algoritmia (Nadarajah ja Kotz [2008], s. 102 – 103). Parametriestimaatit saadaan kiinnitettyjen vapausasteiden tapauksessa iteratiivisesti päivittämällä kaavoja

$$\begin{aligned} s_i^{(k)} &= (x_i - \mu^{(k)})^T S_{(k)}^{-1} (x_i - \mu^{(k)}) \\ w_i^{(k)} &= \frac{\nu + p}{\nu + s_i^{(k)}} \\ \mu^{(k+1)} &= \frac{\sum_{i=1}^n w_i^{(k)} x_i}{\sum_{i=1}^n w_i^{(k)}} \\ S_{(k+1)} &= \frac{1}{n} \sum_{i=1}^n w_i^{(k)} (x_i - \mu^{(k+1)}) (x_i - \mu^{(k+1)})^T. \end{aligned}$$

Edellä μ on jakauman odotusarvo, S skaalamatriisi, ν vapausasteiden lukumäärä ja x_i on i . p -ulotteinen havainto. Vapausasteet $\nu = 3$ valittiin, jotta jakauman hännät olisivat paksut, eikä priorijakauma olisi liian informatiivinen. Estimoinnin tuloksena saatiin

$$(\widehat{\log e_{tgn}}, \widehat{\log k_{tgn}}) = (2.579, -1.942) \quad \text{ja} \quad \hat{S} = \begin{bmatrix} 0.218 & 0.157 \\ 0.157 & 1.148 \end{bmatrix}.$$

Potilaiden parametriestimaatit ja näihin sovitetun t-jakauman tasa-arvokäyrät on visualisoitu kuvaan 11.

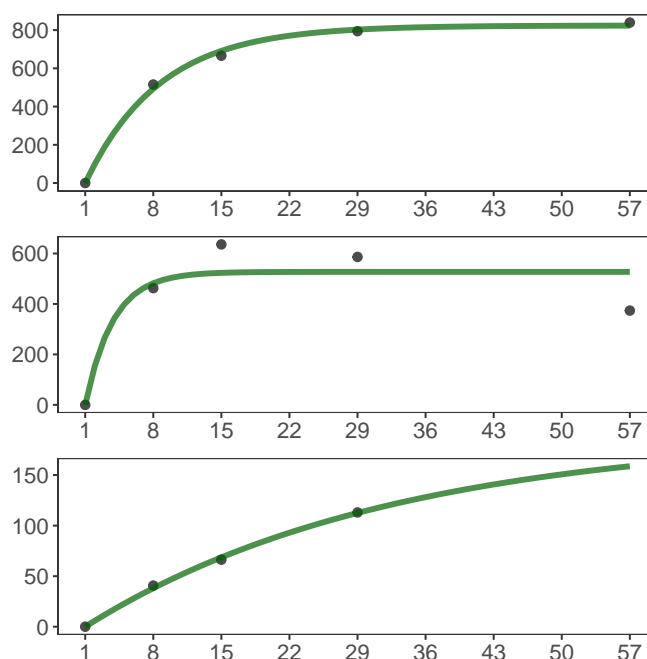
Alustavassa mallinnuksessa kuitenkin havaittiin, että parametrien e_{tgn} ja k_{tgn} identifioituminen on heikkoa parametrin E_{C50} ollessa estimoitavana parametrina. Tämän vuoksi

¹² Hollantilaisten miesten ja naisten keskimääräiset pituudet ovat 182.5 ja 168.7 senttimetriä [NCD Risk Factor Collaboration, 2016].

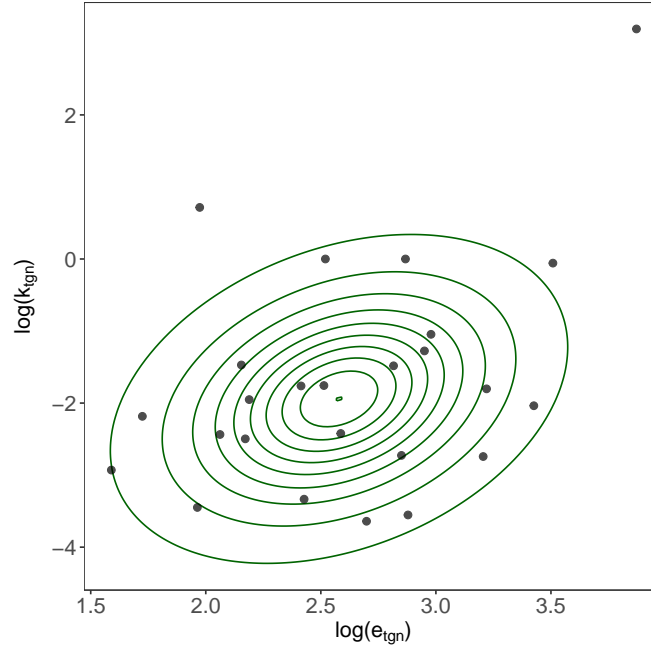
Taulukko 10: Bayesilaisen epälineaarisen tila-avaruusmallin parametrien priorijakaumat tai kiinnitetyt arvot.

Parametri	Priorijakauma tai arvo
$\log(L_0)$	$N(2.115, 0.244)$
$\log(t_{MTT})$	$N(1.151, 0.587)$
$\log(\beta_{vol})$	$N(0, 2)$
$\log(\theta_{ou})$	$N(0, 2)$
$\log(\sigma_{ou})$	$N(0, 2)$
$\log(\sigma_{ip})$	$N(0, 2)$
$\log(E_{C50})$	$N(6.750, 2.168)$
$\log(e_{tgn})$	2.579
$\log(k_{tgn})$	-1.942
γ	0.17
E_{max}	1
σ_{leuk}	0.057
σ_{crp}	0.1

parametrien e_{tgn} ja k_{tgn} arvot kiinnitettiin (logaritmiskaalassa) tämän työn mallinnuksessa EM-algoritmillä estimoituun piste-estimaattiin. Sovitettua t-jakaumaa on kuitenkin mahdollista käyttää priorijakaumana parametreille e_{tgn} ja k_{tgn} mallissa, jossa estimoitavat parametrit on valittu toisin. Parametrin E_{C50} valintaa estimoitavaksi parametriksi parametrien e_{tgn} ja k_{tgn} sijaan perustelee se, että sen estimoimiseen ei ollut käytössä ulkoista aineistoa. Lisäksi parametrin E_{C50} arvon tiedetään vaihtelevan potilaiden välillä 6MP:n yksilöllisen farmakokinetiikan vuoksi.



Kuva 10: Yksinkertaistetun farmakokineettisen mallin sovitte kolmen Derijks ym. [2004] aineiston potilaan TGN_{RBC} -konsentraatioihin. Vaaka-akselilla on hoitopäivä. Pystyakselilla on TGN_{RBC} $8 \cdot 10^8$ punasolua kohti. Mustat pisteet kuvaavat potilaan havaintoja ja vihreä käyrä on mallin sovitte. Potilaiden TGN_{RBC} :n tasapainokonsentraatioissa sekä tasapainokonsentraation saavuttamiseen kuluva ajassa esiintyy vaihtelua, mikä vaikuttaa järkevältä 6MP:n yksilöllisen farmakokinetiikan perusteella.



Kuva 11: Parametrien e_{tgn} ja k_{tgn} potilaskohtaisten parametriestimaattien logaritmeihin sovitettu kaksiulotteinen t -jakauma kolmella vapausasteella. Mustat pisteet ovat potilaille estimoidut parametriestimaatit. Vihreät ellipsit kuvaavat jakauman tasa-arvokäyriä. Tasa-arvokäyrät rajaavat sisäpuolelleen todennäköisyydet 0.7489, 0.6194, 0.5146, 0.4232, 0.3405, 0.2643, 0.1930, 0.1257, 0.0617 ja 0.0004.

6.3 Estimointi

Bayesilaisen epälineaarisen tila-avaruusmallin parametrien estimointi tehtiin käyttäen kohdassa 5.3 esiteltyä adaptiivista MCMC-algoritmia ja laajennettua Kalman-suodinta (*extended Kalman filter*) (Durbin ja Koopman [2012], s. 227 – 228). Laajennettu Kalman-suodin on rekursiivinen algoritmi ns. suodatettujen tilaestimaattien ja suodatettujen kovarianssiestimaattien laskemiseksi epälineaarisen tila-avaruusmallin tapauksessa. Epälineaarisen tila-avaruusmallin (13) merkinnöin suodatetut tilaestimaatit tarkoittavat ehdollisia odotusarvoja

$$\mu_{t|t} = E[\alpha_t | Y_{1:t} = y_{1:t}],$$

ja suodatetut kovarianssiestimaatit ehdollisia kovariansseja

$$\Sigma_{t|t} = Cov[\alpha_t | Y_{1:t} = y_{1:t}],$$

kun $t = 1, \dots, n$ ja parametrivektori θ on kiinnitetty. Edellä y on moniulotteinen aikasarja, jonka pituus on n aikapistettä ja merkintä $y_{1:t}$ tarkoittaa aikasarjan y havaintoja ajanhetkiltä $\{1, \dots, t\}$. Laajennettu Kalman-suodin perustuu Kalman-suotimeen (*Kalman filter*), jonka avulla suodatetut tila- ja kovarianssiestimaatit voidaan laskea lineaarisen tila-avaruusmallin tapauksessa (Durbin ja Koopman [2012], s. 82 – 86). Laajennetussa Kalman-suotimessa epälineaarisen tila-avaruusmallin havainto- ja tilayhtälö linearisoidaan Taylor-approksimaatiolla ja Kalman-suodinta hyödynnetään näin saatuaan lineaariseen tila-avaruusmalliin (Durbin ja Koopman [2012], s. 226 – 227). Taylor-approksimaatiosta johtuen laajennetulla Kalman-suotimella tehty tilastollinen päättely on likimääräistä. Al-

goritmin sivutuotteena saadaan likimääräinen parametrivektoria θ vastaava uskottavuusfunktion arvo (Durbin ja Koopman [2012], s. 170 – 171).

Laajennettu Kalman-suodin muodostuu epälineaarisen tila-avaruusmallin (13) merkinnöin toisiaan seuraavista kaavoista (Durbin ja Koopman [2012], s. 228):

$$\begin{aligned}
\mu_{t|t-1} &= T_t^{(\theta)}(\mu_{t-1|t-1}) \\
\Sigma_{t|t-1} &= \dot{T}_t^{(\theta)}\Sigma_{t-1|t-1}(\dot{T}_t^{(\theta)})' + R_t^{(\theta)}(\mu_{t-1|t-1})R_t^{(\theta)}(\mu_{t-1|t-1})' \\
v_t &= y_t - Z_t^{(\theta)}(\mu_{t|t-1}) \\
F_t &= \dot{Z}_t^{(\theta)}\Sigma_{t|t-1}(\dot{Z}_t^{(\theta)})' + H_t^{(\theta)}(\mu_{t|t-1})H_t^{(\theta)}(\mu_{t|t-1})' \\
\mu_{t|t} &= \mu_{t|t-1} + \Sigma_{t|t-1}(\dot{Z}_t^{(\theta)})'F_t^{-1}v_t \\
\Sigma_{t|t} &= \Sigma_{t|t-1} - \Sigma_{t|t-1}(\dot{Z}_t^{(\theta)})'F_t^{-1}\dot{Z}_t^{(\theta)}\Sigma_{t|t-1}.
\end{aligned} \tag{22}$$

Edellä $\dot{T}_t^{(\theta)}$ on funktion $T_t^{(\theta)}$ Jacobin matriisi pisteessä $\mu_{t-1|t-1}$ ja $\dot{Z}_t^{(\theta)}$ on funktion $Z_t^{(\theta)}$ Jacobin matriisi pisteessä $\mu_{t|t-1}$. Laajennetussa Kalman-suotimessa lasketaan ensin ajanhetken t tilavektorin odotusarvo $\mu_{t|t-1}$ ja kovarianssin estimaatti $\Sigma_{t|t-1}$. Tätä vaihetta kutsutaan algoritmin ennustevaiheeksi (*prediction step*) (Durbin ja Koopman [2012], s. 84). Ennustevaiheessa saadut tilaennusteet vastaavat tietämystä hetken t tilavektorin jakaumasta ehdolla edelliset $t - 1$ havaintoa. Hetken t havainto y_t sisältää kuitenkin lisäinformaatiota hetken t tilavektorin jakaumasta, josta havainto mallin mukaan riippuu. Algoritmin päivitysvaiheessa (*update step*) havainnon y_t lisäinformaatio saman ajanhetken tilavektorista otetaan huomioon laskemalla suodatetut tila- ja kovarianssiestimaatit $\mu_{t|t}$ ja $\Sigma_{t|t}$ (Durbin ja Koopman [2012], s. 83). Laajennetun Kalman-suotimen ideana on ennuste- ja päivitysvaiheita vuorottelemalla korjata sekventiaalisesti aiempien aikapisteidien perusteella tehtyä tilavektorin ennustetta havainnon sisältämällä lisäinformaatiolla.

Tilavektorista riippumaton puuttuvuus aikasarjan y havainnoissa voidaan huomioida (laajennetussa) Kalman-suotimessa. Hetken t havainnon y_t puuttuessa kokonaan algoritmin päivitysvaihe jätetään tekemättä ja asetetaan $\mu_{t|t} = \mu_{t|t-1}$ ja $\Sigma_{t|t} = \Sigma_{t|t-1}$ (Durbin ja Koopman [2012], s. 110 – 112). Jos hetken y_t havainto puuttuu osittain, eli se on havaittu vain joidenkin alkioden osalta, havainto y_t ja matriisit $Z_t(\mu_{t|t-1})$, $\dot{Z}_t^{(\theta)}$ sekä $H_t(\mu_{t|t-1})$ korvataan kaavassa (22) sopivalla matriisilla W_t kerrotuilla vastineillaan. Matriisi W_t valitaan siten, että tulo $y_t^* = W_t y_t$ on vektori, joka sisältää vain ne havaintovektorin y_t alkioita, jotka on havaittu kyseisellä aikapisteellä.

Log-uskottavuusfunktion arvo parametrivektorille θ saadaan kaavasta (Durbin ja Koopman [2012], s. 171):

$$l(\theta) = -\frac{np}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^n (\log |F_t| + v_t' F_t^{-1} v_t), \tag{23}$$

jossa F_t ja v_t lasketaan laajennetussa Kalman-suotimessa (22) ja p on havaintovektorin dimensio sekä n havaintojen lukumäärä kuten edellä. Osittain puuttuvien havaintojen tapauksessa p voi riippua ajanhetkestä t , jolloin kaavan (23) termi np korvataan summalla

$$\sum_{t=1}^n p_t,$$

jossa p_t on havaintovektorin y_t havaituista alkioista muodostetun vektorin dimensio (Durbin ja Koopman [2012], s. 113 – 114).

Normeeraamaton posteriorijakauma on bayesilaisen epälineaarisen tila-avaruusmallin tapauksessa marginaaliposteriorijakauma

$$p(\theta|y) \propto p(y|\theta)p(\theta),$$

jonka uskottavuusfunktion $p(y|\theta)$ logaritmin arvo saadaan kaavasta (23). Laajennettu Kalman-suodin lasketaan siten jokaisella parametrivektorin θ ehdotuksella, joka MCMC-algoritmissa simuloidaan. Priorijakauman $p(\theta)$ tiheysfunktion arvo saadaan tulona taulukon 10 jakaumien tiheysfunktioiden arvoista samaan tapaan kuin bayesilaisen differentiaaliyhtälömallin tapauksessa. Bayesilaisen epälineaarisen tila-avaruusmallin sovittamisessa hyödynnettiin R-kielen *bssm*-pakettia, jolla adaptiivista MCMC:tä voidaan soveltaa suureen joukkoon erilaisia tila-avaruusmalleja [Helske ja Vihola, 2018].

6.4 Konvergenssi

Bayesilainen epälineaarinen tila-avaruusmalli sovitettiin kutakin potilasta kohden simuloimalla kolme 100000 iteraation MCMC-ketjua. MCMC-simuloinnin tehostamiseksi kutakin MCMC-ketjua ja potilasta kohden suoritettiin normeeraamattoman posteriorijakauman numeerinen optimointi alkuarvovektorista alkaen ennen MCMC-simuloinnin aloittamista. Alkuarvovektorit numeeriselle optimoinnille valittiin satunnaisesti moniulotteisesta normaalijakaumasta, jonka komponenttien varianssit olivat kunkin parametrin suuruusluokkaan nähden suuria. Optimointia jatkettiin, kunnes posteriorijakaumalle löydettiin lokaali moodi. Kussakin optimoinnissa löydettyä moodia käytettiin parametrivektorin θ alkuarvona θ_0 . Löydetyssä moodissa laskettiin lisäksi sitä vastaavan Hessen matriisin käänteismatriisi, jonka Cholesky-hajotelmaa käytettiin edelleen adaptiivisen MCMC-algoritmin ehdotusjakauman Cholesky-hajotelman alkuarvona S_0 . Kuvatun menettelyn vuoksi ketjujen lämmitystä vähennettiin 10000 iteraatioon. Ehdotusjakauman kovarianssimatriisin adaptointia jatkettiin koko simuloinnin ajan. Simulointi kesti noin 15 tuntia samalla laitteistolla, jolla simulointi tehtiin bayesilaisen differentiaaliyhtälömallin tapauksessa.

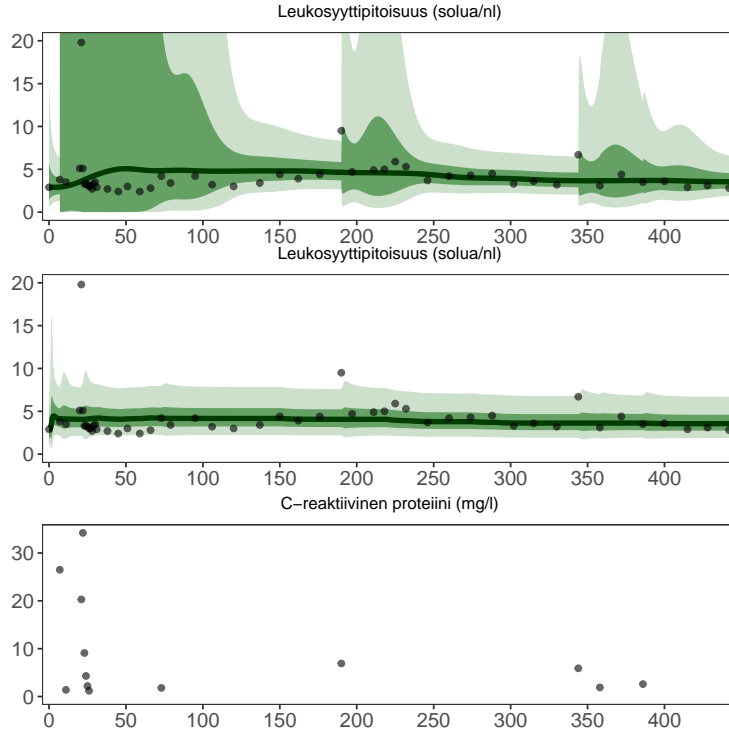
Potilaiden posteriorijakaumien simulointi onnistui hyvin. Potilasta 10 lukuun ottamatta kaikkien potilaiden posteriorijakaumat olivat MCMC-ketjujen perusteella yksihiippuisia. Potilaan 10 MCMC-ketjujen ei oleteta käytetyn simulointialgoritmin vuoksi edustavan hyvää otosta potilaan posteriorijakaumasta.

Potilailla, joilla posteriorijakaumat olivat yksihiippuisia, ylärajat \hat{R}_u olivat kaikille parametreille lähellä arvoa 1, joten MCMC-ketjujen voidaan todeta konvergoineen. Varmuutta konvergenssista lisättiin tutkimalla sattumanvarainen MCMC-ketju kultakin potilaalta graafisesti.

Kaikkien parametrien efektiiviset otoskoot olivat potilaittain yhdistetyissä MCMC-ketjuissa joitakin poikkeuksia lukuun ottamatta suuria (minimi: 462, Q_1 : 4200, Md : 6920, Q_3 : 8936, maksimi: 11605).

6.5 Potilaiden posteriorijakaumat

Potilaan 10 posteriorijakauman kaksihiippuisuus näkyi siten, että potilaan MCMC-ketjuista yksi simuloi näytteitä pelkästään jakauman ensimmäisen moodin ympäristöstä kahden muun ketjun näytteistäessä jakauman toisen moodin ympäristöä. Taulukko 11 esittää



Kuva 12: Bayesilaisen tila-avaruusmallin mukainen leukosyyttipitoisuuden jakauma potilaan 10 posteriorijakauman kahden eri moodin perusteella. Kahdessa ylimmässä kuvassa on samat havainnot, mutta mallin havaintojen mukainen jakauma on laskettu posteriorijakauman kahden eri moodin ympäristöstä poimituilla näytteillä. Ylimmässä kuvassa tulehdusarvo vaikuttaa merkittävästi mallin mukaisen havaintojen jakauman todennäköisyysväleihin, keskimmaisessa ei. Mallin mukaisen havaintojen jakauman 50 ja 90 %:n todennäköisyysvälit ulottuvat ylimmässä kuvassa kuvan ulkopuolelle. Kuvassa näkyvien todennäköisyysvälien laskenta on selitetty kappaleessa 6.6.

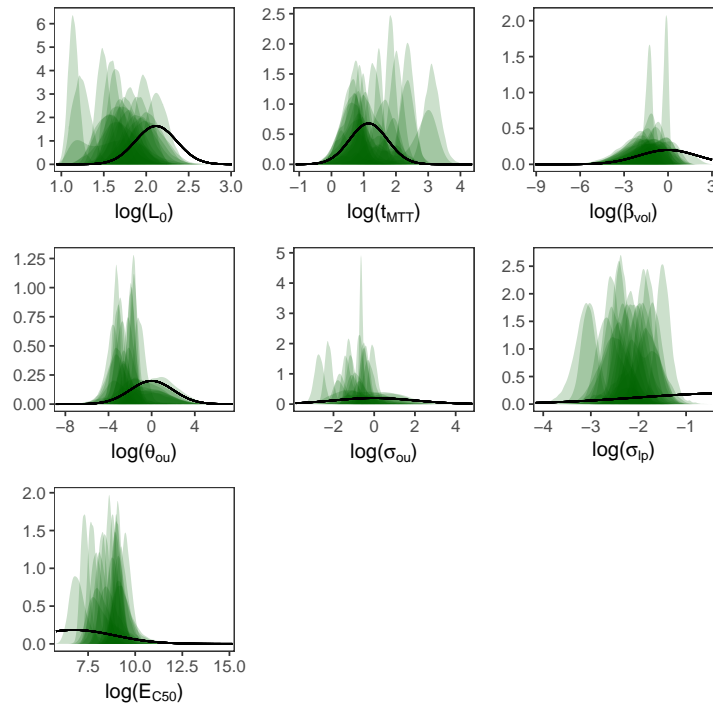
potilaan keskimääräiset parametrien arvot eri moodeja näytteistäneiden MCMC-ketjujen perusteella. Kuva 12 esittelee mallin mukaisen havaintojen jakauman eri ketjuista poimituilla näytteillä. Eri moodien ympäristössä olevat parametrien arvot näyttävät kuvaa vaihtoehtoisia selityksiä samalle aineistolle. Ketjun 1 perusteella lasketussa mallin mukaisessa havaintojen jakaumassa tulehdusarvo vaikuttaa merkittävästi leukopoiesissa olevaan epävarmuuteen, kun taas ketjun 2 tapauksessa tulehdusarvolla ei juuri ole vaikutusta. Ero selittyy erityisesti parametrien t_{MTT} ja β_{vol} marginaalijakaumissa olevien erojen perusteella. Ketjun 2 mukaisessa jakaumassa parametrit t_{MTT} ja β_{vol} saavat suurempia arvoja kuin ketjun 1 mukaisessa jakaumassa, jonka vuoksi potilaan leukopoiesi toimii hitaammin ja tilamuuttujan V vaikutus leukopoiesiin epävarmuuteen on suurempi.

Osa mallin parametreista oli voimakkaasti korreloituneita niiden potilaiden posteriorijakaumissa, jotka olivat yksihuippuisia. Erityisesti parametrien EC_{50} ja L_0 välillä havaittiin voimakkaita negatiivisia korrelaatiota (minimi: -0.966, Q_1 : -0.919, Md : -0.870, Q_3 : -0.845, maksimi: -0.762). Myös Ornstein-Uhlenbeck-prosessin parametrit θ_{ou} ja σ_{ou} (minimi: 0.381, Q_1 : 0.818, Md : 0.898, Q_3 : 0.970, maksimi: 0.990) sekä leukopoiesin nopeuteen vaikuttava parametri t_{MTT} ja parametri σ_{lp} (minimi: -0.906, Q_1 : -0.837, Md : -0.731, Q_3 : -0.696, maksimi: -0.186) olivat korreloituneita.

Kuvan 4 tapaan piirretty kuva 13 esittelee potilaskohtaiset marginaaliposteriorijakaumat taulukon 10 mukaisiin priorijakaumiin verrattuna. Kuvan tarkastelussa on mukana kaikki aineiston potilaat potilasta 10 lukuun ottamatta. Käytössä oleva aineisto näyttää

Taulukko 11: Potilaan 10 parametrien keskiarvot eri moodeja näytteistäneiden MCMC-ketjujen perusteella. Suurimmat erot keskiarvoissa havaitaan parametrien t_{MTT} , β_{vol} , σ_{ou} , θ_{ou} ja σ_{lp} välillä. Ketju 3 jätettiin taulukon ulkopuolelle, sillä se näyttöisesti saman moodin ympäristöä kuin ketju 2.

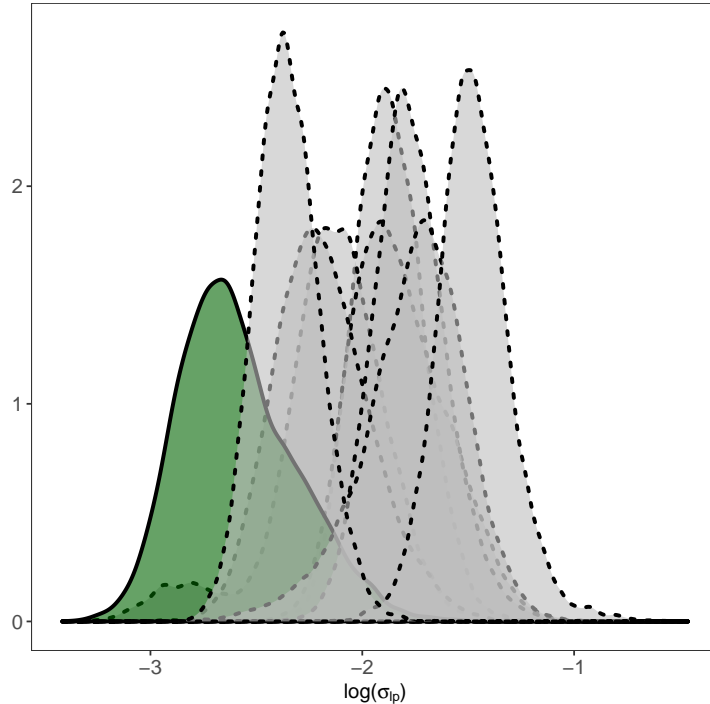
Parametri	Ketju 1	Ketju 2
$\log(t_{MTT})$	2.593	-0.137
$\log(L_0)$	1.625	1.449
$\log(E_{C50})$	9.608	10.381
$\log(\beta_{vol})$	0.477	-2.743
$\log(\theta_{ou})$	-0.917	-1.951
$\log(\sigma_{ou})$	0.508	0.046
$\log(\sigma_{lp})$	-2.624	-1.222



Kuva 13: Bayesilaisen epälineaarisen tila-avaruusmallin parametrien potilaskohtaiset marginaaliposteriorijakaumat mallinnuksessa käytettyyn priorijakaumaan verrattuna.

sisältävän informaatiota kaikista mallin parametreista, sillä marginaaliposteriorijakaumat eroavat käytetyistä priorijakaumista.

Kuva 14 esittää parametrin σ_{lp} marginaaliposteriorijakaumat (logaritmisskaalassa) niiden potilaiden osalta, joille on tehty TPMT-genotyypin määrittäminen. Kuvassa TPMT-heterosygotin potilaan 5 marginaaliposteriorijakauma on merkitty vihreällä kuten aiemmin kuvan 5 ja parametrin σ_{leuk} tapauksessa. Kuvaa 14 (ja 13) tarkastelemalla voi havaita, että potilaan 5 leukopoeesiin liittyvä epävarmuus tilamuuttujan V arvon ollessa 0 on pienempää verrattuna TPMT-villityyppiä oleviin aineiston potilaisiin, joilla TPMT-genotyyppi on tiedossa. Tarkastelu ei ole sama kuin kuvassa 5, sillä tarkasteltava parametri σ_{lp} ei ole bayesilaisessa epälineaarissa tila-avaruusmallissa samassa asemassa kuin bayesilaisen differentiaaliyhtälömallin parametri σ_{leuk} .



Kuva 14: Parametrin σ_p marginaaliposteriorijakaumat niiden potilaiden osalta, joille on tehty TPMT-genotyypin määrittäminen. Potilas 10, jonka posteriorijakauma oli kaksihuippuinen, ei ole kuvan tarkastelussa mukana. Kuva on piirretty samaan tapaan kuin kuva 5.

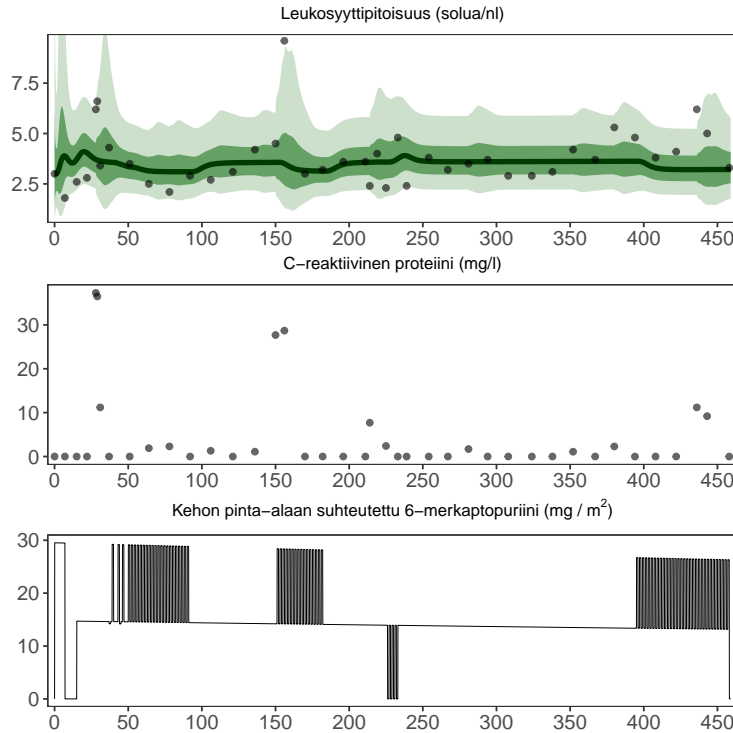
6.6 Mallin mukaiset leukosyyttipitoisuudet

Kuvat 15 – 18 esittelevät bayesilaisen tila-avaruusmallin sovitteet potilaiden 3, 4, 6 ja 5 aineistoissa. Vastaavat kuvat esitettiin potilaista 3, 4 ja 6 kappaleessa 5.6. Samassa kappaleessa esillä ollut potilas 2 jätettiin tarkastelun ulkopuolelle, sillä mallin rakenteen vuoksi kiinnostuksen kohteena on tarkastella erityisesti potilaita, joiden elimistössä on ollut tulehdus hoidon aikana. Annetun 6MP:n sijaan kuviin on piirretty aiemmasta poiketen potilaan kehon pinta-alan arvioon suhteutettu 6MP. Kuviin on piirretty mallin mukaisen havaintojen jakauman kvartiiliväli sekä keskimäinen 90 % jakaumasta kullakin ajanhetkellä kuten aiemminkin.

Mallin mukainen havaintojen jakauma on laskettu poimimalla 500 näytteen otos potilaan parametrien marginaaliposteriorijakaumasta ja laskemalla laajennettu Kalman-suodin uudelleen kutakin poimittua näytettä kohden siten, että leukosyyttipitoisuudet on asetettu puuttuviksi mallin havaintovektorissa. Tämä antaa mahdollisuuden vertailla havaittuja leukosyyttipitoisuuksia mallin mukaiseen leukosyyttipitoisuuden jakaumaan, joka perustuu vain annettuun lääkitykseen, havaittuihin tulehdusarvoihin sekä potilaan parametrien marginaaliposteriorijakauman sisältämään informaatioon mallin käyttäytymisestä. Mallin mukainen leukosyyttipitoisuuden jakauma hetkellä t on normaalijakuumien sekoitusjakauma, jonka komponenttijakaumien odotusarvot ja keskihajonnat ovat

$$\alpha_{L,t|t}^{(\theta_i)} \quad \text{ja} \quad \sqrt{\Sigma_{L,t|t}^{(\theta_i)} + \sigma_{leuk}^2}, \quad (24)$$

tässä järjestyksessä. Yläindeksointi parametrien marginaaliposteriorijakaumasta poimitulla parametrivektorilla θ_i ($i \in \{1, \dots, 500\}$) viittaa edellä parametrivektoria vastaaviin



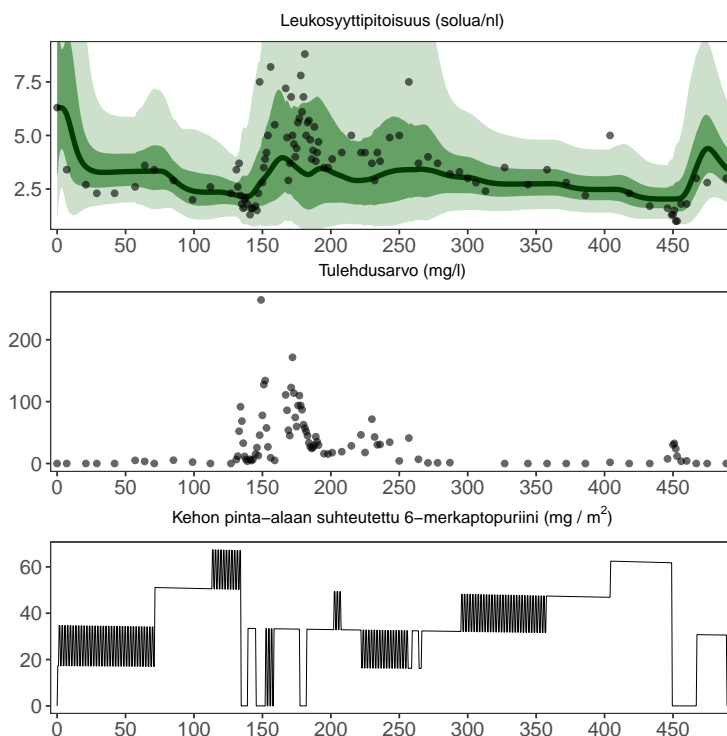
Kuva 15: Bayesilaisen epälineaarisen tila-avaruusmallin sovite potilaan 3 aineistoon. Keskimääräinen 90 % mallin mukaisesta havaintojen jakaumasta ulottuu kuvan ulkopuolelle.

suodatettuihin estimaatteihin ja alaindeksi $L, t|t$ viittaa tilamuuttujaa L vastaavaan suodatetun tila-estimaatin alkioon tai suodatetun kovarianssiestimaatin diagonaalialkioon. Kaavassa (24) tilamuuttujan L suodatettuun varianssiestimaattiin lisätään havaintovarianssi σ_{leuk}^2 , jotta havaintojen epävarmuus tulisi otetuksi huomioon. Kuvissa näkyvät fraktiilit on laskettu numeerisesti kunkin ajanhetken sekoitusjakaumasta.

Kuva 15 esittelee mallin sovituksen potilaan 3 aineistoon. Potilaan 3 hoidon aikana on ollut vain lyhytkestoisia tulehduksia. Noin päivien 25, 150, 225 ja 450 kohdilla koholla oleva tulehdusarvo pullistaa hetkellisesti mallin mukaisen havaintojen jakauman todennäköisyysväljää. Samoihin aikoihin havaitut poikkeavat leukosyyttipitoisuudet näyttävät tämän seurauksena todennäköisemmiltä kuin bayesilaisen differentiaaliyhtälömallin tapauksessa ja sijoittuvat noin päivän 150 mittausta lukuun ottamatta 90 %:n todennäköisyysvälille. Kun tulehdusta ei ole, todennäköisyysvälit ovat kapeampia, eli mallin varmuus leukosyyttipitoisuudesta on suurempi.

Sama ilmiö näkyy myös potilasta 4 koskevassa kuvassa 16. Potilaan hoidon aikana on ollut pitkäaikainen tulehdus noin päivien 125 – 275 välillä. Epävarmuus leukosyyttipitoisuudesta on kuvasta katsottuna erittäin suurta tulehduksen aikana. Päivien 150 – 175 aikainen mallin mukaisen havaintojen jakauman 90 %:n fraktiili näyttää nyt realistisemmalta kuin bayesilaisen differentiaaliyhtälömallin tapauksessa. Jakauma on lisäksi merkittävästi kapeampi päivien 25 – 125 ja päivien 300 – 450 aikana, jolloin potilaan elimistössä ei ole ollut tulehdusta. Näillä väleillä malli näyttää onnistuvan suhteellisen hyvin.

Noin päivän 400 kohdalla samassa kuvassa havaitaan poikkeava havainto, joka on mallin mukaisen havaintojen jakauman 90 %:n fraktiilia suurempi. Poikkeavaan havain-



Kuva 16: Bayesilaisen epälineaarisen tila-avaruusmallin sovite potilaan 4 aineistoon. Keskimääräinen 90 % mallin mukaisesta havaintojen jakaumasta ulottuu kuvan ulkopuolelle.

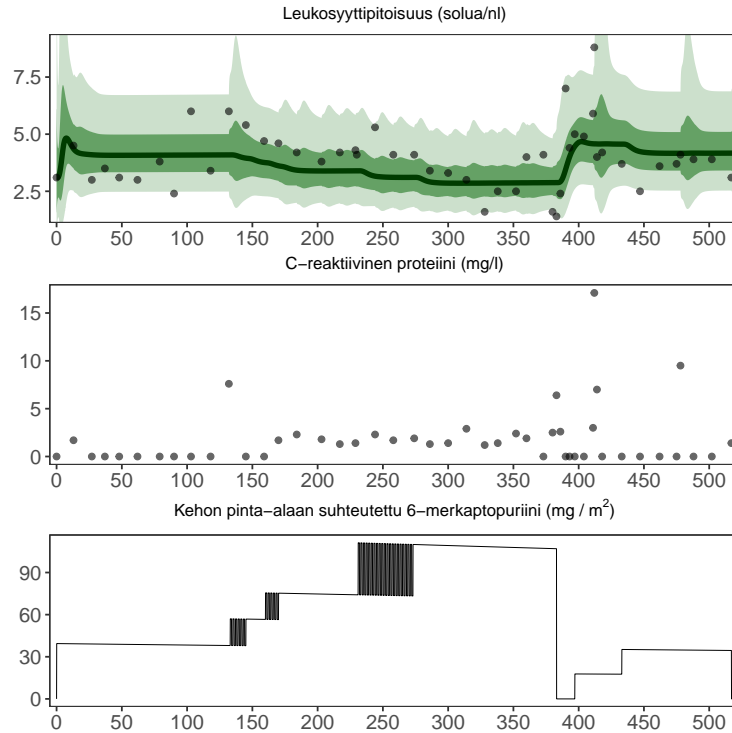
toon on yhteydessä hieman koholla oleva tulehdusarvo¹³, joka on kuitenkin päivien 125 – 275 tulehdusjakson aikana mitattuihin tulehdusarvoihin verrattuna pieni. Tästä huolimatta poikkeava leukosyyttipitoisuus poikkeaa jakauman mediaanista likipitäen yhtä paljon kuin useat potilaan leukosyyttipitoisuuksista päivien 125 – 275 tulehdusjakson aikana. Koska matala tulehdusarvo ei tällä kohtaa selitä leukosyyttipitoisuuden poikkeamaa, on mahdollista, että poikkeaviin leukosyyttipitoisuuksiin liittyy jokin toinen tekijä, joka tulisi huomioida leukopoeesin volatilititeettimallinnuksessa. Koska havaittu tulehdusarvo on pieni, voi poikkeavan leukosyyttipitoisuuden selittää myös virustulehdus, joka ei kappaleeseen 3 viitaten kasvata tulehdusarvoa yhtä paljon kuin bakteeritulehdus. Toisaalta tulehdusarvohavainnoille valittu muunnos tai funktion σ_{wn} valinta voivat myös vaikuttaa havainnon poikkeavuuteen.

Potilaan 4 kuvasta huomataan lisäksi, että mallin mukainen havaintojen jakauma reagoi bayesilaisen differentiaaliyhtälömallin vastaavaa paremmin potilaan lääkityksen muutoksiin. Tämä voidaan havaita erityisesti tarkastelemalla ajanhetkittäisten jakaumien mediaanien kehitystä.

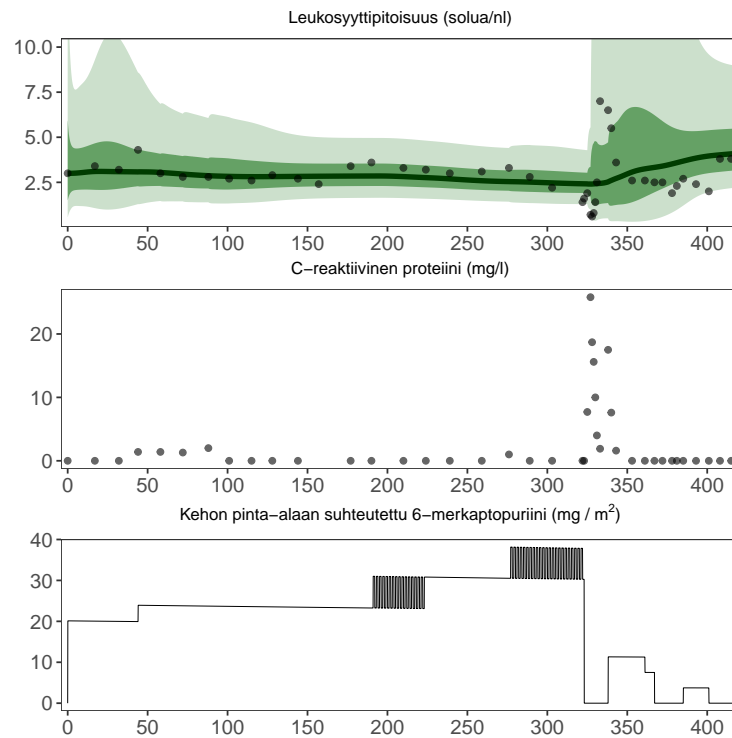
Tämä seikka korostuu erityisen hyvin myös kuvassa 17, jossa lääkitystä on madallettu merkittävästi noin päivän 400 kohdalla. Mallin mukaisen havaintojen jakauman nähdään nyt asettuvan uudelle tasolle, toisin kuin bayesilaisen differentiaaliyhtälömallin tapauksessa, jonka mukainen havaintojen jakauma ei juuri muuttunut tämän potilaan aineistossa lääkityksen keskeyttämisen seurauksena. Paremmasta lääkitykseen reagoinnista huolimatta malli kuitenkin aliennustaa noin päivien 150 – 300 välillä.

Potilasta 5 koskevassa kuvassa 18 parametrille t_{MTT} estimoitu suuri arvo (keskiarvo lineaariskaalassa 15.661 päivää) saa aikaan leveiden todennäköisyysvälien pitkän keston

¹³ Havainnon tarkka arvo on 1.9. Y-akselin skaala on kuvassa leveä.



Kuva 17: Bayesilaisen epälineaarisen tila-avaruusmallin sovite potilaan 6 aineistoon. Keskimäinen 90 % mallin mukaisesta havaintojen jakaumasta ulottuu kuvan ulkopuolelle.



Kuva 18: Bayesilaisen epälineaarisen tila-avaruusmallin sovite potilaan 5 aineistoon. Keskimäinen 90 % mallin mukaisesta havaintojen jakaumasta ulottuu kuvan ulkopuolelle.

hoitohistorian alussa ja lopussa. Hoitohistorian alun leveisiin todennäköisyysväleihin ei liity tulehdusta, vaan todennäköisyysvälit ovat leveät leukopoeiesimallin tilamuuttujien

alkuepävarmuuden suuruudesta johtuen. Hoitohistorian loppupuolella todennäköisyysvälien leventyminen johtuu noin päivän 325 aikaan äkillisesti alkaneesta tulehduksesta. Parametrin t_{MTT} suuresta arvosta johtuen potilaan leukopoeesi toimii hitaasti, jonka vuoksi myös mallin mukainen leukosyyttipitoisuus ja todennäköisyysvälit muuttuvat hitaasti. Tästä johtuen hoitohistorian alussa sekä lopussa havaittavat leveät todennäköisyysvälit kestävät kuvasta katsoen liian pitkään. Tämä on harmillista erityisesti päivänä 325 alkaneen tulehdusjakson osalta, sillä leveiden todennäköisyysvälien rajoittuminen pelkästään tulehdusjaksolle olisi toivottavaa. Vastaavaa ilmiötä ei havaita kuvissa 15 – 17, sillä näiden kuvien potilailla parametrin t_{MTT} arvo on pienempi.

Parametrin t_{MTT} havaittiin aiheuttavan leveiden todennäköisyysvälien pitkän keston simulaatiossa, jossa potilaan muut parametrien arvot asetettiin MAP-estimaattiin, mutta parametrin t_{MTT} arvoa muokattiin pienemmäksi. Simulaation tuloksena oli kuvan 18 kaltainen kuva, jossa pitkään jatkuvia leveitä todennäköisyysvälejä ei havaittu hoitohistorian alkupuolella, ja loppupuolen todennäköisyysvälit olivat leveitä vain tulehdusjakson aikaisina päivinä 325 – 350. Tässä simulaatiossa malli kuitenkin yliennusti päivien 350 – 425 leukosyyttipitoisuudet reilusti. Parametrin t_{MTT} arvolla vaikuttaa olevan suurehko vaikutus mallin mukaiseen havaintopäpävarmuuteen.

Kiinnostavaa kuvassa on lisäksi se, että noin päivän 325 aikaan madallettu lääkitys ei tulehdustilan jälkeen juurikaan vaikuta aineiston havaittuihin leukosyyttipitoisuuksiin. Kenties tämän vuoksi lääkeannoksen muutokseen reagoiva malli yliennustaa leukosyyttipitoisuuksia.

Kuvien 15 – 18 perusteella vaikuttaa siltä, että leukosyyttipitoisuuteen liittyvän epävarmuuden arviointi on bayesilaiseen differentiaaliyhtälömalliin verrattuna realistisempaa, kun tulehdusarvo on huomioitu mallissa. Lisäksi käytetty 6MP:n farmakokineettinen malli vaikuttaa toimivan paremmin kuin farmakokineettinen malli (7). Kuva 18 antaa kuitenkin viitteitä sille, että mallin rakenteessa on kenties parantamisen varaa. Lisäksi mallin mukaisessa havaintojen jakaumassa oleva epävarmuus on edelleen melko suurta, vaikka potilaan elimistössä ei olisikaan tulehdusta. Tulehdusarvon rooli leukopoeesin epävarmuuteen yhteydessä olevana muuttujana vaikuttaa kuitenkin järkevältä.

7 Yhteenveto ja pohdintaa

ALL-lapsipotilaiden veren leukosyyttipitoisuuksia mallinnettiin edellä kahdella bayesilaisella mallilla taudin ylläpitovaiheen aikana. Tässä kappaleessa tuodaan yhteen työssä tehdyn mallinnuksen tuloksia, kommentoidaan Jayachandran ym. [2014, 2015] tekemää mallinnusta ja pohditaan bayesilaisen epälineaarisen tila-avaruusmallin kontribuutiota, kehitysideoita sekä hyödyntämistä.

7.1 Bayesilainen differentiaaliyhtälömalli ja Jayachandran ym. [2014, 2015] mallinnus

Tämän työn malleista ensimmäinen oli kappaleessa 5 esitelty bayesilainen differentiaaliyhtälömalli, joka perustui Jayachandran ym. [2014, 2015] 6MP:n farmakokineetiikkaa sekä leukopoeesia kuvaaviin malleihin, jotka esiteltiin kappaleessa 4.2. Bayesilaisessa differentiaaliyhtälömallissa farmakokineettisen mallin (7) tilamuuttujaa X_{tgn} ei havaittu, sillä kappaleessa 3 kuvatusta aineistosta puuttui Jayachandran ym. [2014, 2015] aineistossa

olleet TGN_{RBC} -mittaukset, jonka suhteen Jayachandran ym. [2014] mallinsivat 6MP:n solumyrkyllisen vaikutuksen leukopoeesiin. Mittausten puuttuminen loi mallinnushaasteen, koska 6MP:n farmakokineettista mallia ei täten voitu sovittaa eikä tutkia erillään leukopoeesimallista.

Jayachandran ym. [2014] oletusta siitä, että TGN_{RBC} -konsentraatio kuvaa hyvin 6MP:n myelosuppressiivisen vaikutuksen, voidaan pitää kyseenalaisena, sillä tutkimusten mukaan TGN_{RBC} :n määrällä on vain heikko yhteys TGN_{DNA} :n määrään, joka kappaleeseen 2 viitaten saa aikaan 6MP:n soluille myrkylliset vaikutukset [Schmiegelow ym., 2014]. Tämän työn mallinnuksessa TGN_{RBC} -mittauksia ei käytetty, mutta mallinnus nojautui kuitenkin osin parametriestimaatteihin ja priorijakaumiin, jotka oli estimoitu käyttäen TGN_{RBC} -mittauksia sisältänyttä aineistoa.

Bayesilaisen differentiaaliyhtälömallin parametrien k_{cm} ja ρ välillä oli korkea korrelaatio usean potilaan posteriorijakaumassa. Lisäksi osalla potilaista posteriorijakauma oli kaksihuippuinen siten, että erot huippujen välillä näkyivät lähinnä edellä mainittujen parametrien marginaaliposteriorijakaumissa. Parametri k_{cm} vaikutti mallissa karkeasti ottaen TGN :n muodostumisnopeuteen ja parametri ρ takaisinkytkennän voimakkuuteen. Näin ollen havaitut korkeat korrelaatiot sekä posteriorijakaumien kaksihuippuisuudet kertonevat siitä, että malli kykenee selittämään leukosyyttipitoisuuden vaihtelun joko muodostuvan TGN :n tai takaisinkytkennän voimakkuuden avulla. Tätä tulkintaa tukee kuva 3, jossa potilaan 15 aineistoon sovitetun mallin keskimääräisen käytöksen nähtiin olevan samanlaista potilaan kaksihuippuisen posteriorijakauman eri moodien ympäristöstä poimittujen näytteiden perusteella. Mallissa vaikuttaa olevan identifioituvuusongelma, joka todennäköisesti ratkeaisi, jos käytössä oleva aineisto sisältäisi farmakokineettisen mallin vastemuuttujana olevat TGN -mittaukset.

Bayesilaisessa differentiaaliyhtälömallissa potilaan lääkityksen muutokset eivät juuri vaikuttaneet mallin mukaisiin leukosyyttipitoisuuksiin. Syynä tälle voi olla kappaleessa 4.2.1 kuvailussa populaatiomallinnuksessa estimoitu parametrien k_{tr} arvo 0.1207, jota mallissa käytettiin. Arvo ei ole linjassa taulukon 3 kanssa, jossa parametrien t_{MTT} arvot olivat välillä 3.73 – 5.62 päivää. Kappaleeseen 6.2 viitaten parametrien t_{MTT} on lisäksi mainittu kirjallisuudessa vaihtelevan välillä 2 – 6.6 päivää. Kaavan (2) perusteella arvo $k_{tr} = 0.1207$ vastaa parametrien t_{MTT} arvo, joka on noin 33 päivää, mikä ei vaikuta edellä mainittujen tietojen perusteella järkevältä. Myöskään taulukon 6 potilaskohtaisesti estimoidut parametrien k_{tr} arvot eivät ole linjassa edellä mainittujen tulosten kanssa. Mahdollinen syy eroavalle parametrien arvoille voi olla esimerkiksi yksikkövirhe Jayachandran ym. [2014] artikkelissa. Poikkeavasta parametrien arvosta seuraava pitkä leukopoeesivaihe vaikuttanee malliin niin, että yhteyttä annetun lääkityksen ja myelosuppressiivisuuden välille ei löydy.

Myös muut Jayachandran ym. [2014, 2015] artikkelien parametrien arvot herättävät kysymyksiä. Kuten kappaleessa 5.1 mainittiin, 6MP:n yksikkönä ei farmakokineettisen mallin (7) tapaan voitu käyttää pikomoolia. Syy tälle voi olla se, että yksikkömuunnosvakion v_{cm} arvo, jota artikkelissa ei mainittu, olikin Jayachandran ym. [2015] mallinnuksessa jokin muu kuin Jayachandran ym. [2014] artikkelissa mainittu 1. Tämän yksityiskohdan lisäksi molemmissa artikkeleissa käytössä olevaa aineistoa hyödynnetään estimoinnissa kahteen kertaan, joka vaikuttanee ilmoitettuihin parametriestimaatteihin.

Edellä tehdyistä huomioista riippumatta bayesilaisen differentiaaliyhtälömallin havaittiin toimivan kohtalaisesti sellaisissa hoitohistorioissa, joissa tulehdukset olivat vähäisiä. Jayachandran ym. [2014, 2015] käyttämät aineistot sisälsivät vain potilaita, joiden hoito

oli onnistunut tasaisella lääkityksellä ilman merkittäviä komplikaatioita, joten tämä vaikuttaa loogiselta. Toisaalta mallin mukaisen havaintojen jakauman todennäköisyysvälit olivat kuitenkin tällaisillakin potilailla esimerkiksi kuvista 6 ja 7 katsottuna melko leveitä. Syynä tälle voi olla esimerkiksi TGN-mittausten puute tai se, että malli ei huomioi muita ilmiön kannalta oleellisia tekijöitä. Osittain mallinnuksen onnistumiseen vaikuttanee käytettyjen priorijakaumien ja parametristimaattien kautta myös se, että Jayachandran ym. [2014, 2015] aineistot sisälsivät aikuisia IBD-potilaita eivätkä ALL-lapsipotilaita.

Bayesilainen differentiaaliyhtälömalli ei toiminut hyvin hoitohistorioissa, joissa tulehduksia oli paljon. Tämä johtuu epäilemättä siitä, että malli ei huomioinut potilaan hoidon aikaisia tulehduksia. On epäselvää, miksi tulehduksia ei huomioitu Jayachandran ym. [2014, 2015] mallinnuksessa. Eräs syy tälle voi olla se, että käytetyt aineistot olivat huomattavasti lyhyempiä kuin tässä työssä käytetyt hoitohistoriat. Kenties siis seurannan pituudesta johtuen tulehduksia ei ole ehtinyt ilmaantua yhtä paljon, eikä muuttujaa siten ole pidetty mallin kannalta oleellisena.

7.2 Pohdintaa bayesilaisesta epälineaarista tila-avaruusmallista

7.2.1 Mallin kontribuutio ja vertailu bayesilaiseen differentiaaliyhtälömalliin

Tulehdusten huomiointi sekä bayesilaisen differentiaaliyhtälömallin mukaisten leukosyyttipitoisuuksien heikko reagointi 6MP-lääkityksen muutoksiin olivat tärkeimmät motivaatiot kappaleessa 6 esitetyn bayesilaisen epälineaarisen tila-avaruusmallin [(14), (15)] kehittämiselle. Malli kehitti leukopoeesimallia (1) lisäämällä siihen tilaepävarmuuden, jonka suuruutta mallinnettiin potilaan tulehdusarvoon yhteydessä olevan tilamuuttujan V suhteen. Tilamuuttuja V noudatti mallissa odotusarvoon hakeutuvaa Ornstein-Uhlenbeck-prosessia. Tulehdusten huomioinnin lisäksi malli yksinkertaisti 6MP:n farmakokinetiikan mallintamista farmakokineettiseen malliin (7) verrattuna.

Esitelty malli sovelsi ensimmäistä kertaa tila-avaruusmalleja ALL:n ylläpitovaiheen leukosyyttipitoisuuksien mallintamiseen. Tila-avaruusmallin käyttäminen mallinnusmenetelmänä sallii differentiaaliyhtälöryhmiä dynaamisemmän mallin, sillä stokastisuus voidaan liittää kaavaan (13) viitaten sekä mallin tila- että havaintoyhtälöön. Mallissa selittäjänä muuttujana ollutta tulehdusarvoa ei ole myöskään työn lähdemateriaalin perusteella aikaisemmin huomioitu ilmiön mallintamisessa. Tulehdusarvon sisällyttäminen malliin on käytetyn aineiston perusteella askel kohti realistisempaa leukosyyttipitoisuuden mallintamista, sillä lähes kaikilla aineiston potilailla on ollut tulehdus hoidon aikana. Tämän perusteella hoidon aikaiset tulehdukset vaikuttavat melko yleisiltä, joten ne on syytä ottaa huomioon ilmiön mallintamisessa.

Kuvien 15 – 18 perusteella mallin mukaiset arviot leukosyyttipitoisuuden epävarmuudesta ovat bayesilaiseen differentiaaliyhtälömalliin verrattuna realistisempia, koska potilaan tulehduksesta johtuva epävarmuus on huomioitu. Tämä on erityisesti nähtävissä kuvasta 16, jossa potilaan tulehdusjakson aikainen leukosyyttipitoisuuden epävarmuus on huomattavasti suurempaa kuin tulehdusjakson ulkopuolella. Malli reagoi myös paremmin 6MP-lääkityksen muutoksiin kuin bayesilainen differentiaaliyhtälömalli. Eron mallien välillä selittänee leukopoeesimallin (1) käyttäminen sekä Friberg ym. [2002] mallinnustulosten ja kirjallisuuden perusteella määritetty parametrin t_{MTT} priorijakauma, joka asettaa parametrin k_{tr} järkevämmälle skaalalle kuin Jayachandran ym. [2014] leukopoeesimallissa.

Myös kappaleessa 6.1.1 esitetty yksinkertaistettu farmakokineettinen malli voi vaikuttaa eron lääkkeen vaikutuksessa mallien välillä. Kuvan 10 perusteella malli sopii hyvin ulkoiseen TGN_{RBC}-aineistoon, jonka avulla parametreille e_{tgn} ja k_{tgn} estimoitiin priorijakauma kappaleessa 6.1. Yksinkertaistettu farmakokineettinen malli korjaa myös farmakokineettisen mallin (7) parametrien k_{cm} ja k_{me} välillä olevan todennäköisen identifioituvuusongelman. Kuten kappaleessa 4.2.2 mainittiin, Jayachandran ym. [2015] raportoivat korkean korrelaation, 0.96, näiden parametrien välillä. Karkeasti ottaen parametrit vaikuttivat mallissa TGN_{RBC}:n muodostumis- ja poistumisnopeuteen. Havaittu korkea korrelaatio johtuu todennäköisesti siitä, että TGN_{RBC}-aineisto voidaan selittää joko suurilla tai pienillä TGN_{RBC}:n muodostumis- ja poistumisnopeuksilla. Differentiaaliyhtälössä (16) parametrit e_{tgn} ja k_{tgn} ovat karkeasti ottaen parametreja k_{cm} ja k_{me} vastaavassa asemassa, mutta niiden välille ei kuitenkaan muodostu vastaavaa korrelaatiota yksinkertaistettua farmakokineettista mallia estimoidessa, sillä TGN:n poistumisnopeuteen vaikuttava parametri k_{tgn} on sisällytetty differentiaaliyhtälössä olevaan funktioon ϕ . Differentiaaliyhtälön (16) versiossa, jossa näin ei ollut, myös parametrien e_{tgn} ja k_{tgn} välille muodostui (logaritmiskaalassa) korkea korrelaatio.

Kaksihuippuinen posteriorijakauma havaittiin bayesilaisen epälineaarisen tila-avaruusmallin tapauksessa vain yhdellä potilaalla¹⁴, mikä voi kertoa siitä, että mallin identifioituvuus on eräiden parametrien välillä olleista korrelaatioista huolimatta parempi kuin bayesilaisen differentiaaliyhtälömallin. Myös mallinnustulosten graafiset tarkastelut näyttävät mielekkäämmiltä kuin työn ensimmäisen mallin tapauksessa, vaikka epävarmuus mallin mukaisessa havaintojen jakaumassa on edelleen suurehkoa.

7.2.2 Mallin kehitysideoita

Uudelleenparametrointi ja TGN-mittaukset

Parametrien t_{MTT} ja σ_{lp} välillä havaittujen korkeiden korrelaatioiden sekä kuvaan 18 liittyvien kommenttien perusteella parametrit olisi kenties hyvä suhteuttaa toisiinsa, joka saattaisi poistaa parametrien välisen korrelaation. Nykyisessä mallissa epävarmuus leukosyyttipitoisuudesta ikään kuin kertyy leukopoeesin ollessa hidas, koska parametreja ei ole suhteutettu toisiinsa. Myös parametrien σ_{ou} ja θ_{ou} välinen korrelaatio olisi kenties mahdollista poistaa toisenlaisella parametroinnilla. Korrelaatio johtuu todennäköisesti siitä, että odotusarvoon (nolla) hakeutuvan Ornstein-Uhlenbeck-prosessin stationaarinen jakauma on

$$N\left(0, \frac{\sigma_{ou}^2}{2\theta_{ou}}\right), \quad (25)$$

joka saadaan tarkastelemalla raja-arvoa $\Delta t_k \rightarrow \infty$ kaavassa (19). Jakaumalle saadaan sama varianssi useilla eri parametrien σ_{ou} ja θ_{ou} valinnoilla, josta korrelaatio johtunee. Parametrien E_{C50} ja L_0 välinen korrelaatio voidaan perustella samaan tapaan kuin bayesilaisen differentiaaliyhtälömallin ja parametrien k_{cm} sekä ρ tapauksessa, koska parametrit e_{tgn} ja k_{tgn} olivat mallissa kiinnitettyinä. On siten todennäköistä, että TGN-mittaukset parantaisivat myös bayesilaisen epälineaarisen tila-avaruusmallin identifioituvuutta.

¹⁴ Potilaan hoitohistoriassa on noin päivän 25 kohdalla poikkeava leukosyyttimittaus, 17.91 solua/nl, jolla voi olla vaikutusta tähän (kuva 18).

Paras 6MP:n myelosuppressiivisesta vaikutuksesta kertova TGN-mittaus olisi kappaleeseen 2.3 viitaten todennäköisesti TGN_{DNA} . TGN_{DNA} -mittauksia sisältäviä aineistoja ei kuitenkaan ole julkisesti saatavilla, eikä 6MP:n farmakokineetiikkaa TGN_{DNA} :ksi ole tämän työn lähdemateriaalin perusteella toistaiseksi mallinnettu kirjallisuudessa. TGN_{RBC} -mittauksia hyödyntävä 6MP:n farmakokineettinen malli saattaa kuitenkin olla hyvä lähtökohta TGN_{DNA} :ta hyödyntävän farmakokineettisen mallin rakentamiselle.

Selittäjien lisääminen malliin

TGN-mittausten avulla tämän työn mallinnusongelma olisi mahdollista jakaa 6MP:n farmakokineettisesta mallista ja leukopoeiesimallista koostuviin pienempiin palasiin, mikä helpottaisi mallikokonaisuuden rakentamista ja tutkimista sekä selittäjien lisäämistä malliin. Kappaleessa 3 kuvaillun aineiston muuttujista potilaan MTX-lääkitys ja TPMT-aktiivisuus ovat ehdokkaita selittäjiksi malliin. Lisäksi virustulehduksesta kertova tai virus- ja bakteeritulehdukset toisistaan erotteleva muuttuja saattaisi olla leukopoeiesin volatiliteettimallinnuksessa hyödyllinen lisämuuttuja, sillä virus- ja bakteeritulehdusten yhteys leukosyyttipitoisuuden epävarmuuteen voi olla kuvaan 16 liittyvien huomioiden valossa erilaista.

MTX-lääkityksellä ja potilaan TPMT-aktiivisuudella on kappaleeseen 2 viitaten yhteys 6MP-lääkityksestä seuraavan solumyrkyllisyyden kanssa, jonka perusteella muuttujat voitaisiin asettaa vaikuttamaan yksinkertaistettuun farmakokineettiseen malliin. MTX-lääkityksen vaikutusta 6MP:n farmakokineetiikkaan voi kuitenkin olla vaikea estimoida esimerkiksi tämän työn aineistosta, jossa 6MP- ja MTX-lääkityksen suuruutta on säädetty leukosyyttipitoisuuden perusteella samanaikaisesti. TPMT-aktiivisuuden vaikuttaminen parametriin e_{tgn} saattaisi kuitenkin olla mielekäs laajennus mallille, sillä esimerkiksi Jayachandran ym. [2015] farmakokineettisessä mallissa TPMT-aktiivisuuden vaikuttaminen parametriin k_{cm} havaittiin mallin kannalta hyödylliseksi. TPMT-aktiivisuudet olivat kuitenkin artikkelin aineistossa punasoluista tehtyjä fenotyypimittauksia, jollaisia tämän työn aineistossa ei ollut.

Kappaleeseen 2.3 viitaten TPMT-fenotyyppi lienee TPMT-genotyyppiä parempi selittäjä, sillä TPMT-aktiivisuudessa on havaittu olevan eroja myös TPMT-genotyyppien sisällä. On kuitenkin mahdollista, että bayesilaisen epälineaarisen tila-avaruusmallin tarkkuudessa on jotakin systematiikkaa myös TPMT-genotyypin mukaan. Tässä työssä TPMT-genotyypin yhteyttä hajontaparametrin σ_{lp} potilaskohtaisiin posteriorijakaumiin tarkasteltiin kuvan 14 avulla. Kuvassa TPMT-heterotsygootin potilaan 5 hajontaparametrin marginaaliposteriorijakauma poikkesi graafisesti tarkasteltuna useimpien TPMT-villityyppejä olevien potilaiden vastaavasta. Tiedot potilaiden TPMT-genotyypeistä olivat aineistossa kuitenkin puutteellisia, joten kyse voi olla myös pelkästä sattumasta.

Tilamuuttujan V vaikutus

Potilaan tulehdusarvoa heijastelevaa tilamuuttujaa V hyödynnettiin tässä työssä leukopoeiesin volatiliteettimallinnuksessa, mutta tilamuuttuja voitaisiin asettaa vaikuttamaan muuhunkin kuin leukopoeiesin epävarmuuteen. [Fliedner ym., 1964] tutkimuksessa kehittyvien neutrofilien verenkiertoon ilmaantuminen tapahtui nopeammin, jos potilaalla oli bakteeritulehdus. Tämän perusteella olisi mielenkiintoista muokata mallia niin, että tilamuuttuja V liitettäisiin parametrin k_{tr} (tai t_{MTT}) arvoon siten, että parametri saisi

suuremman (tai pienemmän) arvon silloin, kun potilaan tulehdusarvo on koholla. Tällöin potilaan leukopoeiesi toimisi mallin mukaan nopeammin tulehduksen aikana, ja malli saattaisi kyetä selittämään paremmin joitakin tulehdustilojen aikana havaittuja leukosyyttipitoisuuksia. Edellä mainittu tutkimus on kuitenkin vanha, ja siinä käytetty aineisto on pieni, joten ilmiö ei välttämättä ole todellinen.

Myös volatiliteettimallinnusta on mahdollista kehittää. Malli, jossa tilamuuttuja V voisi vaikuttaa leukopoeiesin epävarmuuteen eri tavoin lyhyellä ja pidemmällä aikavälillä niin, että tilamuuttujille $S - L$ sallittaisiin omat volatiliteettikertoimet, voisi esimerkiksi olla mielenkiintoinen. Tilamuuttujan V voisi asettaa vaikuttamaan vaihtoehtoisesti myös mallin havaintoyhtälön hajontaparametriin σ_{leuk} . Tällainen malli painottaisi estimoinnissa havaintoja, joita mitattaessa ei ole ollut tulehdusta. Toisaalta solupitoisuuksien odotusarvoinen käytös saattaisi olla tällaisessa mallissa tulehdustilojen aikana epärealistista. Tässä työssä tilamuuttujan V vaikutus leukopoeiesin tilamuuttujien keskihajonnan logaritmiin oli lineaarinen. Vaikutus voisi kuitenkin olla myös epälineaarinen, ja riippua yhdestä tai useammasta parametrasta, jotka estimoidaan.

Tilamuuttujien alkujakaumat

Tilamuuttujan M alkujakauman asettaminen tehtiin mallissa olettaen, että potilaan elimistössä ei hoidon aluksi ole leukosyyttipitoisuuteen vaikuttavaa TGN:ää (tai muita solumyrkkyjä). Oletus ei ole realistinen, sillä kappaleeseen 2.1 viitaten ylläpitovaihe on viimeinen ALL:n hoitovaihe, jota ennen potilas on jo saanut solunsalpaajahoitoa. Siten on mahdollista, että hoidon alussa potilaan elimistössä on aineita, joilla on vaikutus potilaan leukosyyttipitoisuuteen. Tämän työn mallinnuksen yhteydessä tutkittiin edellisten tietojen valossa realistisempaa malliversiota, jossa potilaan elimistössä olisi sallittu olevan hoidon aluksi leukosyyttipitoisuuden kasvua hillitsevä parametrin M_{init} suuruinen TGN-konsentraatio. Käytössä olevalla aineistolla mallissa oli kuitenkin esiteltyjä malleja vakavampia identifioituvuusongelmia, jotka kuitenkin saattaisivat ratketa aineistolla, jossa on TGN-mittaukset.

Leukopoeiesin oletettiin tässä työssä olevan hoidon aluksi likipitäen tasapainotilassa, mitä voitaneen pitää epärealistisena oletuksena samasta syystä kuin tilamuuttujan M tapauksessa. Eräs vaihtoehto tilamuuttujien $S - L$ alkujakaumien määrittämiseksi olisi estimoida alkujakaumien parametrit, mutta tämä lisää mallin estimoitavien parametrien lukumäärää. Työn alustavan mallinnuksen aikana tutkittiin malliversiota, jossa leukopoeiesin tilamuuttujien alkujakaumien odotusarvot määräytyivät kolmen parametrin L_{init} , C_{init} ja S_{init} perusteella. Työssä käytettyyn oletukseen leukopoeiesin tasapainotilasta päädyttiin kuitenkin mallin yksinkertaistamisen vuoksi.

Esitettyä tilamuuttujan V alkujakauman valintaa järkevämpi valinta olisi kenties nollaan hakeutuvan Ornstein-Uhlenbeck-prosessin stationaarinen jakauma (25).

Prediktiiviset vertailut

Tässä työssä ei tarkasteltu bayesilaisen epälineaarisen tila-avaruusmallin tuottamia ennusteita, vaan mallin sovitteita ja käytöstä kokonaisissa potilasaineistoissa. Koska työn taustalla on ennustemallin kehittäminen, tulisi erilaisia malleja tutkia ja vertailla tehtyjen tarkastelujen lisäksi myös niiden ennustekyvyn perusteella. Tarkastelut, jossa malli sovitaan koko aineistoon, eivät kerro suoraan mallin ennustekyvystä. Luonteva tapa mal-

lien ennustekyvyn vertailuun olisi tämän työn hoitohistorioiden tapauksessa esimerkiksi aikasarjoille sopiva ristiinvalidointi (Hyndman ja Athanasopoulos [2014], s. 51 – 52).

Ristiinvalidoinnissa parhaan mallin valinta malliehdokkaista M_1, M_2, \dots, M_k tehdään sovittamalla kukin malleista M_i jokaisen potilaan j ($j = 1, 2, \dots, p$) hoitohistoriaan d_j useita kertoja siten, että mallin sovittamiseen käytettävää aineiston määrää kasvatetaan vähitellen. Jokaisen mallin sovituksen jälkeen lasketaan mallin posteriorijakauman avulla leukosyyttipitoisuuden prediktiivinen jakauma esimerkiksi seuraavalle yhdelle tai kahdelle ajanhetkelle, joilta aineistossa on havainto. Mallin mukaisen prediktiivisen jakauman ja saman ajanhetken havainnon avulla voidaan tämän jälkeen laskea ennustetarkkuutta kuvaavia tunnuslukuja, kuten jakauman moodin ja havainnon välinen keskineliövirhe tai havainnon rajaama häntätodennäköisyys ajanhetken prediktiivisessä jakaumassa. Mallien sovittamista kasvavalla aineiston määrällä jatketaan, kunnes kaikkien potilaiden aineistot on käyty läpi kaikilla malleilla. Mallien vertailua voidaan lopuksi tehdä laskettujen tunnuslukujen avulla.

Ristiinvalidointi on MCMC-menetelmillä kuitenkin laskennallisesti vaativaa, sillä mallit on sovittettava potilasaineistoihin useita kertoja. Eräs tapa nopeuttaa tätä olisi korvata MCMC-simulointi esimerkiksi MAP-estimoinnilla. Tällöin prediktiivinen jakauma ei kuitenkaan huomioi mallin parametrien epävarmuutta. Tämän korjaamiseksi yksi vaihtoehto olisi käyttää ristiinvalidoinnissa MCMC-menetelmien sijaan posteriorijakauman normaaliapproksimaatiota ja tärkeysotannan avulla tehtävää approksimaation korjausta (Gelman ym. [2014], s. 314 – 316). Tämän työn mallinnuksen taustalla tehdyissä MAP-estimoinneissa kävi ilmi, että posteriorijakauman normaaliapproksimaatio on usean potilaan tapauksessa melko hyvä.

Lisäksi kappaleessa 6.4 kuvattu tapa tehostaa adaptiivista MCMC:tä todettiin hyödylliseksi, sillä optimoinnin avulla adaptiiviselle MCMC:lle saadaan hyvät alkuarvot, eikä pitkää lämmitysjaksoa tarvita. Kenties optimointiin yhdistämällä myös MCMC:tä voitaisiin hyödyntää ristiinvalidoinnissa.

Hierarkkinen mallinnus

Bayesilainen epälineaarinen tila-avaruusmalli sovitettiin yhden potilaan aineistoon kerrallaan. Parempia mallinnustuloksia saataisiin kenties muuntamalla malli hierarkkiseksi malliksi, jossa potilaiden parametrivektorit $\theta_1, \theta_2, \dots, \theta_k$ riippuisivat hyperparametrivektorista ϕ sekä jakaumasta $p(\theta|\phi)$ (Gelman ym. [2014], s. 101, 107). Posteriorijakauma olisi tällaisessa mallissa muotoa

$$p(\theta, \phi|y) \propto p(y|\theta)p(\theta|\phi)p(\phi),$$

jossa y sisältää kaikkien potilaiden havainnot, parametrimatriisi θ kaikkien potilaiden parametrivektorit ja $p(\phi)$ on hyperparametrivektorin ϕ priorijakauma. Hierarkkinen malli luo potilaiden parametrivektorien välille riippuvuusrakenteen, jonka kautta kunkin potilaan mallinnustuloksiin vaikuttaa myös muiden potilaiden hoitohistorioista saatu informaatio. Tällä tavoin saatu lisäinformaatio helpottaisi sellaisten potilaiden parametrien estimointia ja identifiointia, joiden hoitohistorioissa informaatio parametreista on vähäinen. Esimerkiksi potilailla, joiden hoitohistorioissa esiintyy vain lyhytaikaisia tulehduksia, informaatio leukopoeesiin liittyvän volatilitteettimallin estimoimiseksi on vähäistä.

Hierarkkisesta mallinnuksesta saatavaa hyperparametrivektorin ϕ posteriorijakaumaa on mahdollista hyödyntää myös aineiston ulkopuolisen potilaan priorijakauman muodosta-

misessa. Tällaisesta priorijakaumasta olisi hyötyä hoidon alkuvaiheen leukosyyttipitoisuuden ennusteiden tekemisessä, sillä priorijakauman avulla epävarmuus potilaan ensimmäisten hoitopäivien leukosyyttipitoisuudesta saataisiin minimoitua ennen potilasta koskevien mittausten kertymistä.

Hyvistä puolistaan huolimatta hierarkkinen mallinnus kuitenkin lisää estimoitavien parametrien lukumäärää ja monimutkaistaa mallia, joten on mahdollista, että se ei onnistuisi ilman TGN-mittauksia.

Kehittyneemmät estimointimenetelmät

Epälineaarisen tila-avaruusmallin estimoinnissa käytettiin tässä työssä approksimatiivista menetelmää, laajennettua Kalman-suodinta. Estimointi on mahdollista tehdä myös mielivaltaisen tarkasti käyttämällä simulointiin perustuvia menetelmiä, joita kutsutaan partikkelisuodattimiksi (Durbin ja Koopman [2012], s. 276 – 299). Posteriorijakauman simuloinnissa on mahdollista käyttää myös partikkeli-MCMC-menetelmiä [Andrieu ym., 2010].

7.2.3 Mallin hyödyntäminen

Ennustemallin käyttämisen ALL:n ylläpitovaiheen hoidossa voidaan yksinkertaistaen ajatella etenevän niin, että potilaan jokaisella kontrollikäynnillä leukosyyttipitoisuuksille tehdään ennusteita ehdolla mittaukset, joita kuhunkin kontrollikäyntiin mennessä on kertynyt. Lääkäri voi ennusteiden avulla tutkia erilaisten lääkeannostuspäätösten ennustamaa leukosyyttipitoisuutta, ja siten lisätä ymmärrystään hoidon tilanteesta. Mallin avulla on kenties mahdollista myös laskea optimaalinen lääkeannostuspäätös tai kokoelma hyviä lääkeannostuspäätöksiä, jotka ennustavat hoidon tilanteelle sopivia leukosyyttipitoisuuksia. Lääkäri voisi hyödyntää tällaisia mallin ehdottamia lääkeannostuspäätöksiä lopullisen lääkeannostuspäätöksen tekemisessä. Huomattavaa on, että mallin ehdottamien lääkeannostuspäätösten laskeminen on kuitenkin ilmiön mallintamisesta erillinen kysymys, jota ei käsitelty tässä työssä.

Leukosyyttipitoisuuden ennusteet olisivat ALL:n hoidossa hyödyllisiä sen vuoksi, että kappaleeseen 2 viitaten nykyiset hoitoprotokollat perustuvat leukosyyttipitoisuuden käyttämiseen hoidon tehokkuuden surrogaattina. Erään tutkimuksen mukaan paras pääasiallisesti seurattava muuttuja lääkeannostuspäätöksissä olisi kuitenkin veren neutrofilipitoisuus [Schmiegelow ym., 2016]. Tämän työn bayesilaisen epälineaarisen tila-avaruusmallin kaltaista mallia on kuitenkin mahdollista hyödyntää myös tilanteessa, jossa vastemuuttujana on neutrofilipitoisuus, sillä malli perustuu leukopoiesimalliin (1), jota on sovellettu kappaleeseen 4.1 viitaten myös neutrofilipitoisuuksien mallintamiseen.

Viimeaikoina ALL:n hoitoon liittyen on tutkittu TGN_{DNA} :n soveltuvuutta vaihtoehdokseksi kriteeriksi, jonka perusteella hoidon tehokkuutta voitaisiin arvioida ja lääkitystä säätää [Nielsen ym., 2017]. Edellä mainitussa tutkimuksessa korkeamman TGN_{DNA} -konsentraation ja harvinaisemman syövän uusiutumattomuuden välille löydettiin yhteys. Tarvetta vaihtoehdokseksi kriteerille saa aikaan se, että leukosyyttipitoisuuksien tavoitevälit eivät kenties välttämättä ole optimaalinen kriteeri lääkityksen suuruuteen liittyvissä päätöksissä, sillä tutkimuksissa on arvioitu, että potilaan tyypillinen veren leukosyyttipitoisuus ilman sairautta voi vaikuttaa pitoisuuksiin, jotka 6MP-lääkityksellä on mahdollista saavuttaa [Schmiegelow ym., 2014]. Lisäksi leukosyyttipitoisuuksiin vaikuttaa esimerkiksi potilaan ikä, etninen tausta sekä päivän- ja vuodenaajat (Haddy ym. [1999],

Haus ja Smolensky [1999]). Koska havaitun leukosyyttipitoisuuden takana on annetun lääkityksen lisäksi myös muita tekijöitä, on mahdollista, että tulevaisuudessa potilaan TGN_{DNA} -konsentraation seuraaminen tulee leukosyyttipitoisuuden rinnalle tai korvaa sen pääasiallisena lääkeannostuspäätösten tekemisessä käytettävänä kriteerinä. Leukosyyttipitoisuuksien ennustaminen on kuitenkin ALL:n hoidon tulevaisuudessa tästä huolimatta todennäköisesti hyödyllistä, sillä leukosyyttipitoisuudella on myös rooli elimistön riittävän hoidonaikaisen puolustuskyvyn ylläpitämisessä. Toisaalta farmakokineettisen mallin muunnelmaa voitaisiin todennäköisesti hyödyntää myös TGN_{DNA} :n mallintamisessa. Näin ollen parhaimmassa tapauksessa bayesilaisen epälineaarisen tila-avaruusmallin kaltaisella mallilla voitaisiin tulevaisuudessa mallintaa kahta hoitoon liittyvää säätökriteeriä samanaikaisesti.

Myös ALL:n ulkopuolella saattaa olla käyttökohteita tässä työssä kehitetyn bayesilaisen epälineaarisen tila-avaruusmallin kaltaiselle mallille, sillä esimerkiksi IBD-potilaita hoidetaan 6MP:llä. Menetelmällisesti ajateltuna tila-avaruusmallit vaikuttavat lisäksi käytökelpoiselta mallinnusmenetelmältä myös muihin lääkkeisiin perustuvien hoitojen tuke-
misessa, joissa potilaalle yksilöllistetyn lääkeannoksen löytäminen on 6MP-hoidon tapaan tärkeää. Esimerkiksi varfariinin (verenohennuslääke) annostelussa kohdataan samankaltaisia haasteita kuin 6MP:n annostelussa, sillä lääkkeen vaarallisten sivuvaikutusten lisäksi lääkevasteessa on viivettä ja yksilöllistä vaihtelua (Carlson [2008], Heneghan ym. [2010]).

Lähteet

- K. Ahnert and M. Mulansky. Odeint – solving ordinary differential equations in C++. In *AIP Conference Proceedings*, volume 1389, pages 1586–1589. AIP, 2011.
- E. K. Amundsen, P. Urdal, T.-A. Hagve, M. R. Holthe, and C. E. Henriksson. Absolute neutrophil counts from automated hematology instruments are accurate and precise even at very low levels. *American journal of clinical pathology*, 137(6):862–869, 2012.
- C. Andrieu, A. Doucet, and R. Holenstein. Particle Markov Chain Monte Carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(3):269–342, 2010.
- F. M. Balis, J. S. Holcenberg, S. Zimm, D. Tubergen, J. M. Collins, R. F. Murphy, G. S. Gilchrist, D. Hammond, and D. G. Poplack. The effect of methotrexate on the bioavailability of oral 6-mercaptopurine. *Clinical Pharmacology & Therapeutics*, 41(4):384–387, 1987. ISSN 1532-6535. doi: 10.1038/clpt.1987.45. URL <http://dx.doi.org/10.1038/clpt.1987.45>.
- D. Brigo and F. Mercurio. *Interest rate models-theory and practice: with smile, inflation and credit*, volume 2. Springer Science & Business Media, 2006.
- B. Carlson. Declaring war on warfarin misdosing. *Biotechnology healthcare*, 5(2):54, 2008.
- Centers for Disease Control and Prevention. Clinical growth charts, 2017. URL https://www.cdc.gov/growthcharts/clinical_charts.htm.

- S. A. Coulthard, C. Howell, J. Robson, and A. G. Hall. The relationship between thiopurine methyltransferase activity and genotype in blasts from patients with acute leukemia. *Blood*, 92(8):2856–2862, 1998.
- J. Dancey, K. A. Deubelbeiss, L. A. Harker, and C. A. Finch. Neutrophil kinetics in man. *Journal of Clinical Investigation*, 58(3):705, 1976.
- L. J. Derijks, L. P. Gilissen, L. G. Engels, L. P. Bos, P. J. Bus, J. J. Lohman, W. L. Curvers, S. J. van Deventer, D. W. Hommes, and P. M. Hooymans. Pharmacokinetics of 6-mercaptopurine in patients with inflammatory bowel disease: implications for therapy. *Therapeutic drug monitoring*, 26(3):311–318, 2004.
- J. Durbin and S. J. Koopman. *Time series analysis by state space methods*, volume 2. Oxford University Press, 2012.
- R. El Edelbi, S. Lindemalm, and S. Eksborg. Estimation of body surface area in various childhood ages - validation of the Mosteller formula. 101:540–4, 12 2011.
- S. Eskelinen. CRP (P-CRP). 2016. URL http://www.terveyskirjasto.fi/terveyskirjasto/tk.koti?p_artikkeli=snk03052. Elektroninen artikkeli, viitattu 25.9.2017.
- T. Fliedner, E. Cronkite, S. Killmann, and V. Bond. Granulocytopoiesis. II. Emergence and pattern of labeling of neutrophilic granulocytes in humans. *Blood*, 24(6):683–700, 1964.
- J. C. G. Franco. Maximum likelihood estimation of mean reverting processes. *Real Options Practice*, 2003. URL http://www.investmentscience.com/Content/howtoArticles/MLE_for_OR_mean_reverting.pdf.
- M. Fransson. *Towards Individualized Drug Dosage-General Methods and Case Studies*. PhD thesis, Institutionen för datavetenskap, 2007.
- L. E. Friberg, A. Henningsson, H. Maas, L. Nguyen, and M. O. Karlsson. Model of chemotherapy-induced myelosuppression with parameter consistency across drugs. *Journal of clinical oncology*, 20(24):4713–4721, 2002.
- P. S. Gaynon. Mercaptopurine in childhood acute lymphoblastic leukaemia. *The Lancet Oncology*, 18(4):425 – 426, 2017. ISSN 1470-2045. doi: [https://doi.org/10.1016/S1470-2045\(17\)30152-3](https://doi.org/10.1016/S1470-2045(17)30152-3). URL <http://www.sciencedirect.com/science/article/pii/S1470204517301523>.
- A. E. Gelfand and A. F. Smith. Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association*, 85(410):398–409, 1990.
- A. Gelman and D. B. Rubin. Inference from iterative simulation using multiple sequences. *Statistical science*, pages 457–472, 1992.
- A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. *Bayesian data analysis*, volume 3. CRC press, 2014.

- T. B. Haddy, S. R. Rana, and O. Castro. Benign ethnic neutropenia: what is a normal absolute neutrophil count? *The Journal of laboratory and clinical medicine*, 133(1): 15–22, 1999.
- E. Haus and M. H. Smolensky. Biologic rhythms in the immune system. *Chronobiology international*, 16(5):581–622, 1999.
- J. Helske and M. Vihola. *bssm: Bayesian Inference of Non-Linear and Non-Gaussian State Space Models in R*, 2018. URL <https://CRAN.R-project.org/package=bssm/vignettes/bssm.pdf>. R package version 0.1.4.
- C. Heneghan, S. Tyndel, C. Bankhead, Y. Wan, D. Keeling, R. Perera, and A. Ward. Optimal loading dose for the initiation of warfarin: a systematic review. *BMC cardiovascular disorders*, 10(1):18, 2010.
- U. Hindorf, M. Lindqvist, C. Peterson, P. Söderkvist, M. Ström, H. Hjortswang, A. Pousette, and S. Almer. Pharmacogenetics during standardised initiation of thiopurine treatment in inflammatory bowel disease. *Gut*, 55(10):1423–1431, 2006.
- R. J. Hyndman and G. Athanasopoulos. *Forecasting: principles and practice*. OTexts, 2014.
- H. Inaba, M. Greaves, and C. G. Mullighan. Acute lymphoblastic leukaemia. *The Lancet*, 381(9881):1943 – 1955, 2013. ISSN 0140-6736. doi: [https://doi.org/10.1016/S0140-6736\(12\)62187-4](https://doi.org/10.1016/S0140-6736(12)62187-4). URL <http://www.sciencedirect.com/science/article/pii/S0140673612621874>.
- D. Jayachandran, A. E. Rundell, R. E. Hannemann, T. A. Vik, and D. Ramkrishna. Optimal chemotherapy for leukemia: a model-based strategy for individualized treatment. *PloS one*, 9(10):e109623, 2014.
- D. Jayachandran, J. Laínez-Aguirre, A. Rundell, T. Vik, R. Hannemann, G. Reklaitis, and D. Ramkrishna. Model-based individualized treatment of chemotherapeutics: Bayesian population modeling and dose optimization. *PloS one*, 10(7):e0133244, 2015.
- P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*. Stochastic Modelling and Applied Probability. Springer, corrected edition, 1995.
- P. Lafolie, S. Hayder, O. Björk, L. Åhström, J. Liliemark, and C. Peterson. Large interindividual variations in the pharmacokinetics of oral 6-mercaptopurine in maintenance therapy of children with acute leukaemia and non-Hodgkin lymphoma. *Acta Paediatrica*, 75(5):797–803, 1986.
- J. M. Laínez-Aguirre, L. Mockus, S. Orçun, G. Blau, and G. V. Reklaitis. A decomposition strategy for the variational inference of complex systems. *Technometrics*, 58(1):84–94, 2016.
- L. Lennard, J. Lilleyman, J. Van Loon, and R. M. Weinshilboum. Genetic variation in response to 6-mercaptopurine for childhood acute lymphoblastic leukaemia. *The Lancet*, 336(8709):225–229, 1990.

- O. Lohi, J. Kanerva, M. Taskinen, A. Harila-Saari, S. Rounioja, K. Jahnukainen, P. Lähteenmäki, and K. Vettenranta. Lapsuusiän leukemia. *Duodecim; lääketieteellinen aikakauskirja*, 129(9):939–946, 2013.
- B. Lund, A. Åsberg, M. Heyman, J. Kanerva, A. Harila-Saari, H. Hasle, S. Söderhäll, Ó. G. Jónsson, S. Lydersen, K. Schmiegelow, ym. Risk factors for treatment related mortality in childhood acute lymphoblastic leukaemia. *Pediatric blood & cancer*, 56(4): 551–559, 2011.
- B. Meibohm and H. Derendorf. Basic concepts of pharmacokinetic/pharmacodynamic (PK/PD) modelling. *International journal of clinical pharmacology and therapeutics*, 35(10):401–413, 1997.
- L. Michaelis and M. L. Menten. *The kinetics of invertase action*. Universitätsbibliothek Johann Christian Senckenberg, 1913.
- Mosteller. Simplified calculation of body-surface area. *New England Journal of Medicine*, 317(17):1098–1098, 1987. doi: 10.1056/NEJM198710223171717. URL <http://dx.doi.org/10.1056/NEJM198710223171717>. PMID: 3657876.
- S. Nadarajah and S. Kotz. Estimation methods for the multivariate t distribution. *Acta Applicandae Mathematicae*, 102(1):99–118, 2008.
- National Cancer Institute. Childhood acute lymphoblastic leukemia treatment. 2018. URL <https://www.cancer.gov/types/leukemia/patient/child-all-treatment-pdq>. Elektroninen artikkeli, viitattu 23.4.2018.
- NCD Risk Factor Collaboration. A century of trends in adult human height. *Elife*, 5, 2016.
- C. M. Nguyen, M. A. Mendes, and J. D. Ma. Thiopurine methyltransferase (TPMT) genotyping to predict myelosuppression risk. *PLoS currents*, 3, 2011.
- S. N. Nielsen, K. Grell, J. Nersting, J. Abrahamsson, B. Lund, J. Kanerva, Ó. G. Jónsson, G. Vaitkeviciene, K. Pruunsild, L. L. Hjalgrim, ym. DNA-thioguanine nucleotide concentration and relapse-free survival during maintenance therapy of childhood acute lymphoblastic leukaemia (NOPHO ALL-2008): a prospective substudy of a phase 3 trial. *The Lancet Oncology*, 18(4):515–524, 2017.
- NOPHO. NOPHO-ALL 2008 final protocol version 3a. 2011. URL http://data.stolav.no/eqspublic/pasientforlop/docs/doc_32989/attachments/protokoll\%20NOPHO\%202008\%20versjon\%20nov\%202012.pdf. Elektroninen artikkeli, viitattu 24.4.2018.
- PubChem. Mercaptopurine. 2018. URL <https://pubchem.ncbi.nlm.nih.gov/compound/667490>. Elektroninen artikkeli, viitattu 28.2.2018.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2017. URL <https://www.R-project.org/>.
- A. Raes, S. Van Aken, M. Craen, R. Donckerwolcke, and J. V. Walle. A reference frame for blood volume in children and adolescents. *BMC pediatrics*, 6(1):3, 2006.

- G. O. Roberts, A. Gelman, W. R. Gilks, ym. Weak convergence and optimal scaling of random walk Metropolis algorithms. *The annals of applied probability*, 7(1):110–120, 1997.
- W. L. Roberts, R. Sedrick, L. Moulton, A. Spencer, and N. Rifai. Evaluation of four automated high-sensitivity C-reactive protein methods: implications for clinical and epidemiological applications. *Clinical Chemistry*, 46(4):461–468, 2000.
- W. Sandborn. Rational dosing of azathioprine and 6-mercaptopurine. *Gut*, 48(5):591–592, 2001.
- K. Schmiegelow, I. Al-Modhwahi, M. K. Andersen, M. Behrendtz, E. Forestier, H. Hasle, M. Heyman, J. Kristinsson, J. Nersting, R. Nygaard, ym. Methotrexate/6-mercaptopurine maintenance therapy influences the risk of a second malignant neoplasm after childhood acute lymphoblastic leukemia: results from the NOPHO ALL-92 study. *Blood*, 113(24):6077–6084, 2009.
- K. Schmiegelow, S. N. Nielsen, T. L. Frandsen, and J. Nersting. Mercaptopurine/methotrexate maintenance therapy of childhood acute lymphoblastic leukemia: clinical facts and fiction. *Journal of pediatric hematology/oncology*, 36(7):503, 2014.
- K. Schmiegelow, J. Nersting, S. N. Nielsen, M. Heyman, F. Wesenberg, J. Kristinsson, K. Vettenranta, H. Schrøeder, R. Weinshilboum, K. L. Jensen, ym. Maintenance therapy of childhood acute lymphoblastic leukemia revisited—should drug doses be adjusted by white blood cell, neutrophil, or lymphocyte counts? *Pediatric blood & cancer*, 63(12):2104–2111, 2016.
- Synlab. Verenkuva ja leukosyyttien erittelylaskenta. 2017. URL <https://www.synlab.fi/laboratoriokasikirja/tutkimuskuvaukset/verenkuva-leukosyyttien/>. Laboratoriokäsikirja, viitattu 13.11.2017.
- M. Vihola. Robust adaptive Metropolis algorithm with coerced acceptance rate. *Statistics and Computing*, 22(5):997–1008, 2012.
- VSSHP. Leukemia lapsilla. 2017. URL <http://www.vsshp.fi/fi/hoito-ja-tutkimukset/syopa/Sivut/leukemia-lapsilla.aspx>. Elektroninen artikkeli, viitattu 6.2.2018.