

Beyond your data

4PHARMA



Marginaalirakennemallit epidemiologisessa tutkimuksessa

Sisältö

- Johdanto
 - Havainnoivat tutkimukset ja aikariippuvainen sekoittuminen
 - Marginaalirakennemallit (Marginal structural models)
- Marginaalirakennemallit
- Mitä pitäisi tapahtua (SOHO –tutkimus)
- Mitä tapahtui (Virusaineiston analyysi)
- Miksi näin

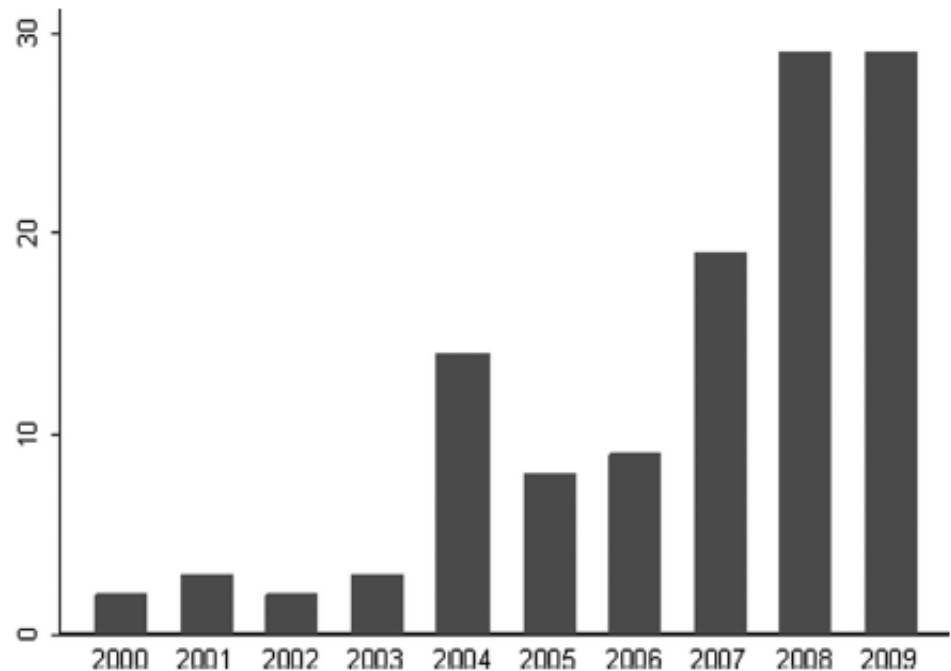
Havainnoivat tutkimukset ja aikariippuvainen sekoittuminen

- Aikariippuvaista sekoittuminen on mahdollista, jos
 - analysoitava aineisto sisältää aikariippuvaisia kovariaatteja, jotka ovat yhteydessä sekä käsittelyyn että vasteeseen.
- Esimerkiksi potilaan yleiskunto voi vaikuttaa potilaan elinaikaan ja annettuun hoitoon.
 - potilaat joiden yleiskunto heikkenee saattavat saada selkeästi enemmän jotain tiettyä lääkitystä.
 - Tällaisessa tilanteessa hyväkin lääkitys saattaa vaikuttaa antavan huonoja tuloksia



Marginaalirakennemallit

- Marginal Structural Models
- Marginaalirakennemallit poistavat aikariippuvaisten sekoittavien tekijöiden aiheuttamaa harhaa painottamalla havaintoja aineistosta muodostettavilla IPT-painoilla (*Inverse probability of treatment*)
- Marginaalirakennemallien käyttö on lisääntynyt tasaisesti vuodesta 1997 lähtien, jolloin James M. Robins esitteli ne ensimmäisen kerran (Robins 1998).



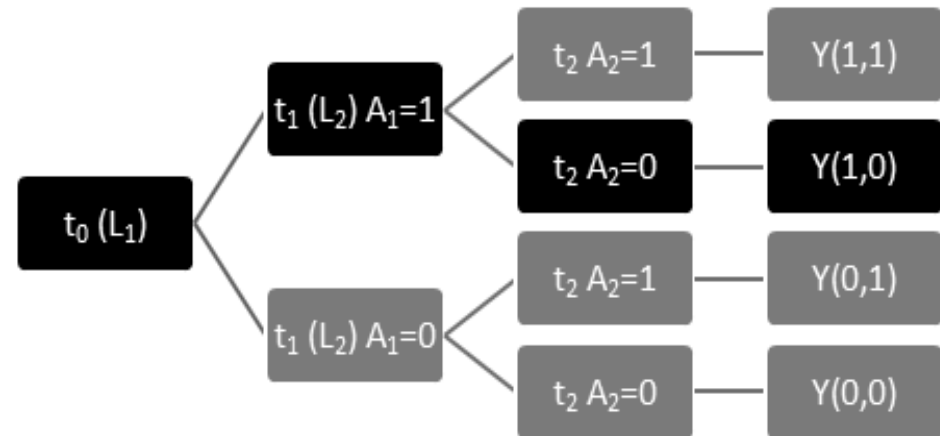
Marginaalirakennemallit

- Menetelmillä, jotka eivät huomioi aikariippuvaista sekoittumista, saadut estimaatit voivat olla voimakkaasti harhaisia
- Marginaalirakennemallissa hyödynnetään IPT-painoja
 - pyritään muodostamaan aineistosta pseudopopulaatio,
 - Satunnaistetun koejärjestelyn tapaan tasapainossa sekoittavien tekijöiden suhteen.
- soveltamalla tyypillistä tilanteeseen sopivaa analyysimenetelmää havaitun aineiston sijasta siitä luotuu pseudopopulaatioon, poistetaan estimateista harhaa

- Korjaa estimateja, mutta kasvattaa luottamusvälejä

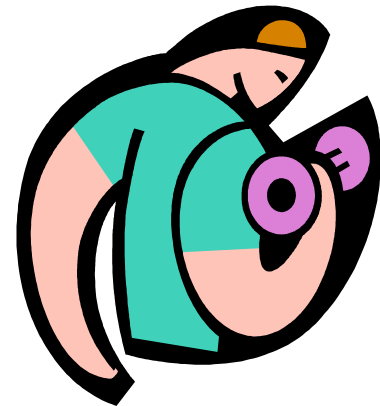
Pseudopopulaatio

- Pseudopopulaation muodostamisessa otetaan huomioon yksilöiden potentiaaliset tapahtumat
- Esimerkkinä kaksivaiheinen aineisto, jota on kerätty alkupisteessä t_0 ja ajanhetkillä t_1 sekä t_2
- Vaste Y havaitaan seurannan lopussa.



IPT-painot

- Marginaalirakennemalleissa IPT-painoilla painotetaan havaintoja niiden käsittelyhistorian todennäköisyyden käänteisluvulla.
- IPT-paino: $w^u = \frac{1}{P(\bar{A}_t|\bar{L}_t)}$
- Vakautettu IPT-paino: $w^s = \frac{P(\bar{A}_t)}{P(\bar{A}_t|\bar{L}_t)}$
- Vakautetut IPT-painot ovat huomattavasti vähemmän vaihtelevia, mutta säilyttävät silti tarvittavat vakauttamattomien IPT-painojen ominaisuudet.



MRM estimointi

- MRM estimointi neljässä vaiheessa
 - (I) Sovita malli käsittelyille
 - (II) Käytä mallia estimoimaan käsittelyiden todennäköisyydet
 - (III) Painota havaintoja IPT-painoilla
 - (IV) Sovita malli vasteelle painotetuilla havainnoilla

Mitä pitäisi tapahtua (SOHO)

- SOHO-tutkimuksessa oli seurattu avohoidossa olevien skitsofreniapotilaiden terveydentilaa.
- Pitkittäinen kolmen vuoden aineisto yli kymmenestä tuhannesta potilaasta
- Tutkimukseen rekrytoidut potilaat olivat joko aloittamassa lääkehoidon tai vaihtamassa nykyistä lääkitystään toiseksi.
- Potilaiden tiedot kerättiin aikapisteissä, jotka olivat lähtötilanne, 3 kuukautta, 6 kuukautta ja tämän jälkeen kuuden kuukauden välein aina kolmeen vuoteen asti lähtötilanteesta.
- Kiinnostavin mittauksen kohde oli lääkärin arvioima skitsofrenian kliininen vakavuus, joka määritettiin arvioimalla mittaushetkellä potilaan oireiden vakavuutta.
 - arvo yhden ja seitsemän välillä, jossa terve saa arvon yksi ja äärimmäisen sairas arvon seitsemän.
- Potilaat saattoivat vaihtaa lääkitystään minä hetkenä hyvänsä ja silti säilyä tutkimuksessa.

Analyysi (SOHO)

- Vasteena käytettiin dikotomista tietoa sairauden remissiotilasta (0 = Remissiokriteerejä ei saavutettu, 1 = Sairaus remissiotilassa).
- Sairauden määriteltiin olevan remissiossa, jos kliininen vakavuustaso oli saanut arvon kolme tai vähemmän kuuden kuukauden ajan.
- Analyysiin oli sisällytetty vain potilaat, joilla oli korkeintaan yksi puuttuva käynti (N=6516).
- Mallissa hyödynnettiin runsaasti mitattuja kovariaatteja lähtötilanteen sosiodemografisista sekä kliinisistä tekijöistä. Näiden lisäksi tekijöiksi asetettiin seurantajakso ja jakson alussa käytetty lääkitys.
- Lopullinen malli muodostettiin poistamalla runsaasti tekijöitä sisältävästä mallista yksittäin tekijöitä, jotka eivät olleet tilastollisesti merkitseviä.
- Tulokset yllättivät: Klotsapiini menestyi odotettua heikommin...



Tulokset (SOHO)

- Lääkärit olivat tietoisia aikaisemmista klotsapiinilla saaduista hyvistä hoitotuloksista
- klotsapiiniä määrättiin useammin vakavasti sairaille potilaille
- Seurauksena, vaikka klotsapiinista oli ollutkin apua, sen vaikutus ei ollut täyttänyt tutkimuksessa määriteltyjä remissiokriteerejä.

Lääke	Perinteiset estimaatit		MRM estimaatit	
	OR	95 % lv	OR	95 % lv
Amisulpride	0.726*	(0.561 - 0.938)	0.835	(0.574 - 1.216)
Clozapine	0.784*	(0.649 - 0.946)	1.038	(0.753 - 1.430)
Depot typical	0.592**	(0.505 - 0.693)	0.771*	(0.622 - 0.956)
Oral typical	0.626**	(0.548 - 0.738)	0.724*	(0.527 - 0.995)
Quetiapine	0.657**	(0.563 - 0.766)	0.729**	(0.578 - 0.919)
Risperidone	0.740**	(0.661 - 0.829)	0.764**	(0.647 - 0.902)
Combination	0.637**	(0.577 - 0.703)	1.012	(0.853 - 1.200)

*0.05 > P > 0.01

**P < 0.01

Referenssinä olanzapine

Tulokset SOHO

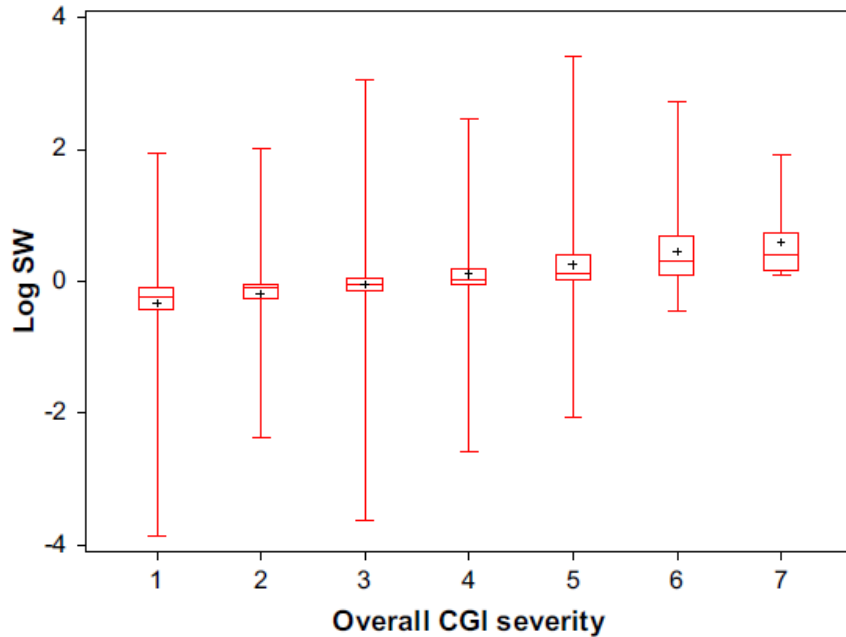


Fig. 1. Boxplots of the logarithm of “stabilized” inverse-probability of treatment weights (SW) for olanzapine by previous overall CGI severity.

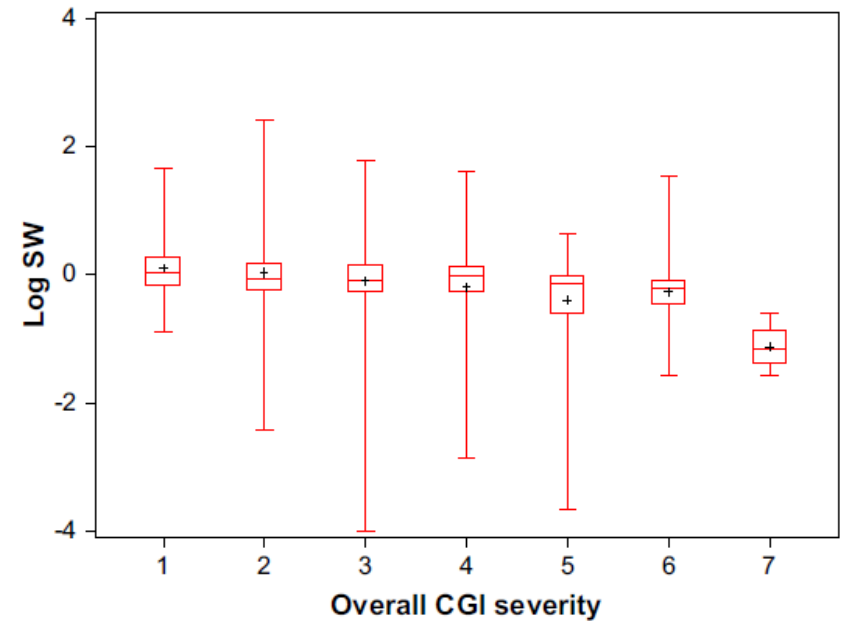


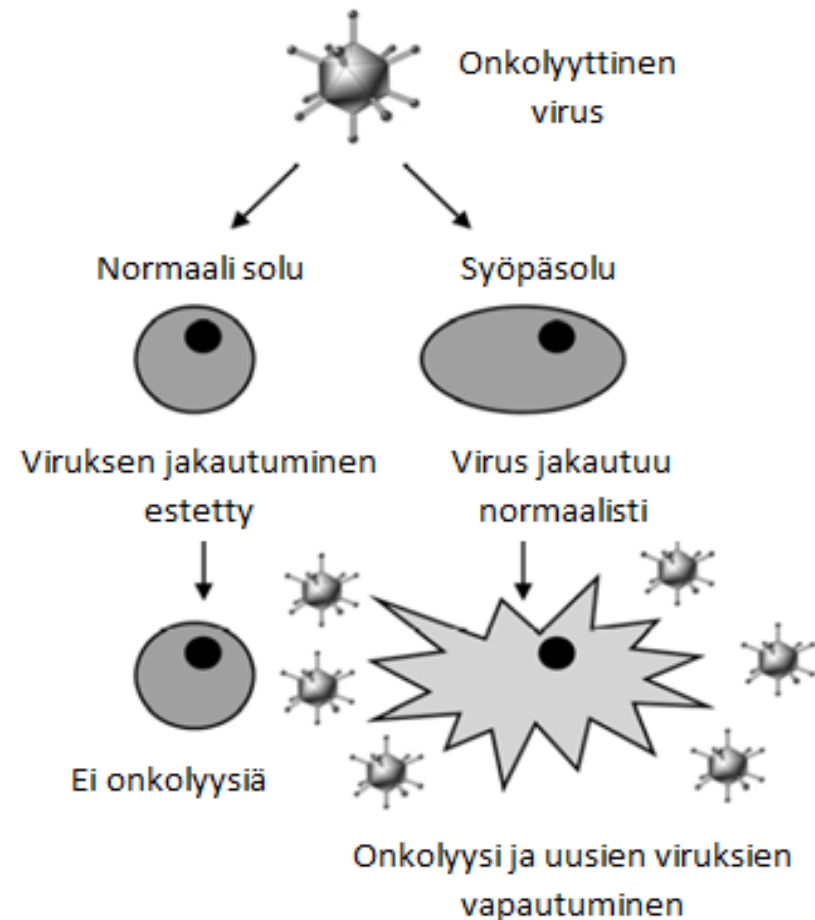
Fig. 2. Boxplots of the logarithm of “stabilized” inverse-probability of treatment weights (SW) for clozapine by previous overall CGI severity.

Mitä tapahtui (Virusaineisto)

- Aineisto
 - Aineistossa oli 274 potilasta.
 - Potilaiden seuranta aloitettiin hetkestä, jolloin he saivat ensimmäisen virushoitonsa
 - Potilailta saadut seuranta-ajat vaihtelivat lyhimmästä kuudesta päivästä aina 1054 päivään asti
 - 213 eventtiä

Onkolyttiset virukset syövän hoidossa

- Serotyyppi
 - AD3, AD5, AD5/3 AD5-RGD
- Immuuniherätteet
 - CD40L, GMCSF
- Lukuisia promoottoreita



Virushoidot

Luokka	Virukset	Hoitojen lkm
AD5/3	AD3-HTERT-E1	32
	AD5/3-COX2L-D24	36
	AD5/3-D24-GMCSF	194
	AD5/3-D24-HNIS	1
	AD5/3-E2F-D24-GMCSF	73
	AD5/3-HTERT-CD40L	118
AD5-RGD	AD5-D24-RGD	16
	AD5-D24-RGD-GMCSF	88
	ICOVIR-7	77
AD5	AD5-D24-GMCSF	113

Toisena mielenkiinnon kohteena pidettiin mahdollista virukseen liitetyn immuuni-herätteen (GMCSF, CD40L) vaikutusta elinaikaan.

Tästä muodostettiin erillistä analyysiä varten dikotominen tekijä.

Taustamuuttujat

	Keskiarvo	Keskihajonta	min.	q1	mediaani	q3	maks.
Ikä	55,5	13,3	3	48	58	64	84
	Luokka	N (%)					
WHO	matala	152 (55,5)					
	korkea	122 (45,5)					
Sukupuoli	mies	116 (42,3)					
	nainen	158 (57,7)					

Aikariippuvaiset muuttujat

- Aikariippuvaisilla tekijöillä pyrittiin huomioimaan mahdollista aikariippuvaista sekoittumista
- kolme erilaista, joilla pyrittiin esittämään potilaan hoitohistoriaa mahdollisimman hyvin.
 - Hoitoaika
 - kuinka kauan potilas on ollut samalla hoidolla määriteltyjen virusluokkien mukaan. Tällä ei siis huomioitu, kuinka monta kertaa potilas on saanut samaa käsittelyä, vaan ainoastaan se, kuinka monta päivää potilas on ollut samantyyppisen virushoidon alla.
 - hoitojen lukumäärä
 - Kertojen määrä, jolloin potilaan hoito oli vaihtunut toisen tyyppiseksi.
 - käsittelyjen lukumäärä
 - yksittäisten hoitokertojen määrää riippumatta siitä, minkälaista hoitoa annettiin.

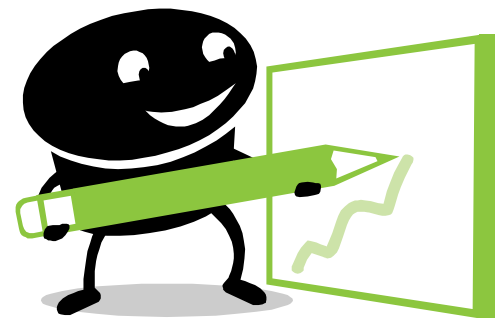
Analyysi

- Analyysimenetelmänä käytettiin diskreettien aikojen elinaika-analyysiä, jossa hyödynnettiin yleistettyjä estimointiyhtälöitä.
- Aineisto jaettiin kymmenen päivän mittaisiin aikaikkunoihin
 - aineiston paloittainen käsittely mahdollisti helpon tavan muodostaa yksittäisille havainnoille käytettävät IPT-painot
- Havaintoja käsiteltiin siten, että jokaisella henkilöllä ajateltiin olevan aikavälillä tietty virushoito aktiivisena. Aikavälin virushoito määritettiin edellisestä hoidosta
 - Kun henkilö sai uuden hoidon, sitä seuraavissa aikaväleissä henkilöllä katsottiin olevan tämä virushoito aktiivinen kunnes henkilö sai taas uuden virushoidon.
- Analyysi suoritettiin GEE-estimoinnilla. Tilastollisia malleja tehtiin kiinnostuksen kohteiden mukaisesti kaksi erilaista. Mallit olivat muuten samanlaisia lukuun ottamatta sitä, että toisessa mallissa yhtenä tekijänä oli virustyyppi ja toisessa tämän tilalla oli dikotominen immuunivastemuuttuja.
- Aikaikkunaa osoittavalle tekijälle tehtiin logaritmi-muunnos mallin sopivuuden parantamiseksi.



Painottamaton malli

- GEE-malli (vasteena elossa/kuolema)
- Kovariaatteina (aikaikkunassa j)
 - Käsittely(j)
 - Sukupuoli
 - WHO
 - Ikä
 - $\text{Log}(\text{Aika}(j))$



Painotettu malli

- Marginaalirakennemallituksessa muodostettiin ensin käsittelymallit, joiden perusteella laskettiin käytetyille havainnoille IPT-painot.
- Jokaista henkilön tilaa eri aikaväleillä käsiteltiin yksittäisinä havaintoina, joille jokaiselle laskettiin oma IPT-paino.
- Lopullinen marginaalirakennemalli, jota hyödynnettiin varsinaisten mielenkiinnon kohteiden estimointiin, vastasi painotettua versiota aikaisemmasta painottamattomasta mallista.

Käsittelymallit

- Vakautetut painot
 - Kaksi mallitusta
 - Ensimmäinen malli sisälsi pelkästään aikariippuvaiset tekijät
 - Toinen malli näiden lisäksi myös lähtötilanteen taustamuuttujat.
- Vakautetut painot määritettiin näistä saatujen estimaattien suhteena. Saadun käsittelyn todennäköisyyttä mallitettiin logistisella regressiolla
- Aikariippuvaisina tekijöinä oli hoitoaika, hoitojen lukumäärä ja käsittelyjen lukumäärä
- Lähtötilanteen taustamuuttujina käytettiin WHO-luokitusta, sukupuolta ja ikää
- Keskiarheet kuitenkin kasvoivat kohtuuttoman suuriksi käytettäessä kaikki tekijöitä ja tarkasteltaessa jokaista aikaväliä.
- Tekijöitä ja tarkasteluaikaa jouduttiin rajaamaan

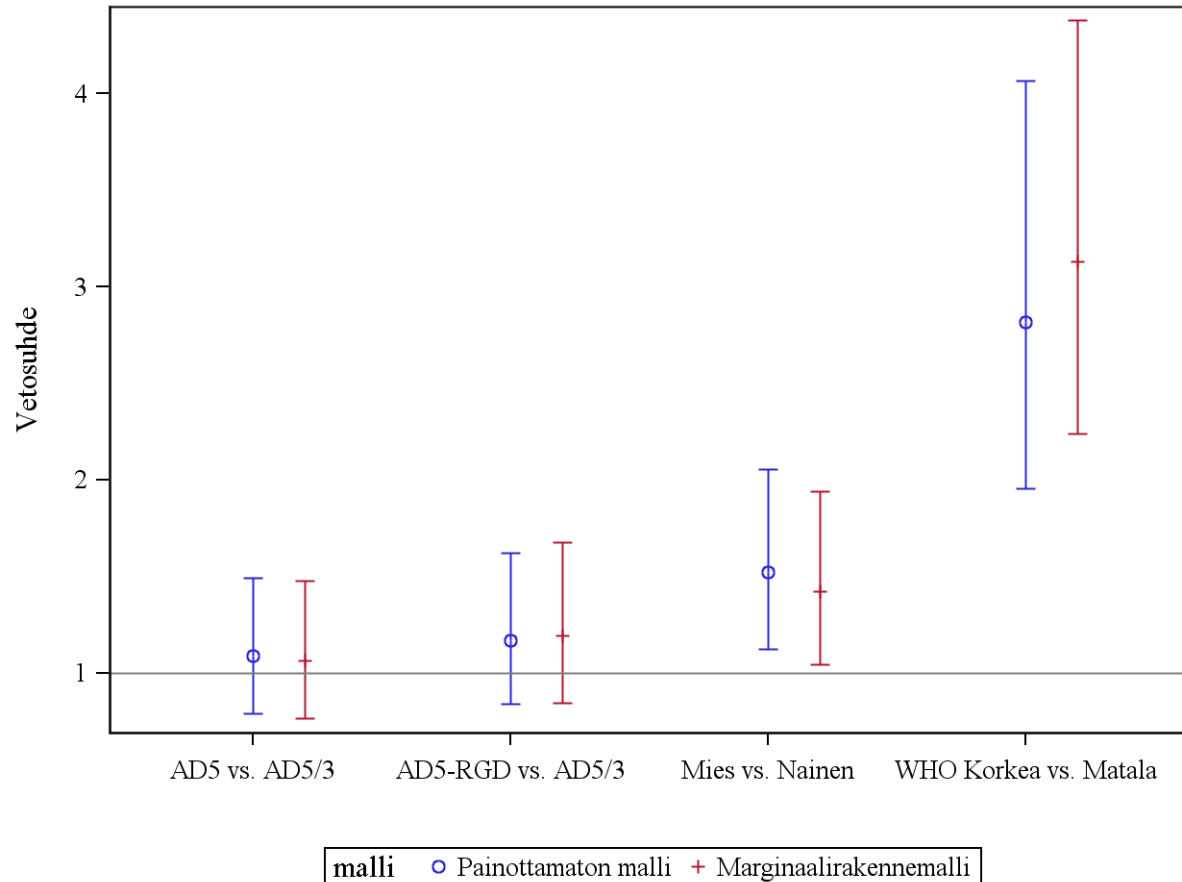
Lopulliset IPT-painot

Malli	Keskiarvo	Keskihajonta	Min.	Q1	Mediaani	Q3	Maks.
Virustyyppi	0.997	0.186	0.339	0.920	0.995	1.051	3.378
Immuuniheräte	1.000	0.133	0.219	0.970	1.000	1.021	3.183

Tulokset ja vertailu

Vertailu	Painottamaton malli		MRM	
	OR	95 % luottamusväli	OR	95 % luottamusväli
Virustyyppi (AD5 vs. AD5/3)	1.087	(0.791, 1.493)	1.062	(0.763, 1.477)
Virustyyppi (AD5-RGD vs. AD5/3)	1.165	(0.839, 1.620)	1.191	(0.846, 1.676)
Sukupuoli (mies vs. nainen)	1.520	(1.123, 2.056)	1.423	(1.044, 1.940)
WHO (Korkea vs. Matala)	2.818	(1.952, 4.066)	3.130	(2.237, 4.380)

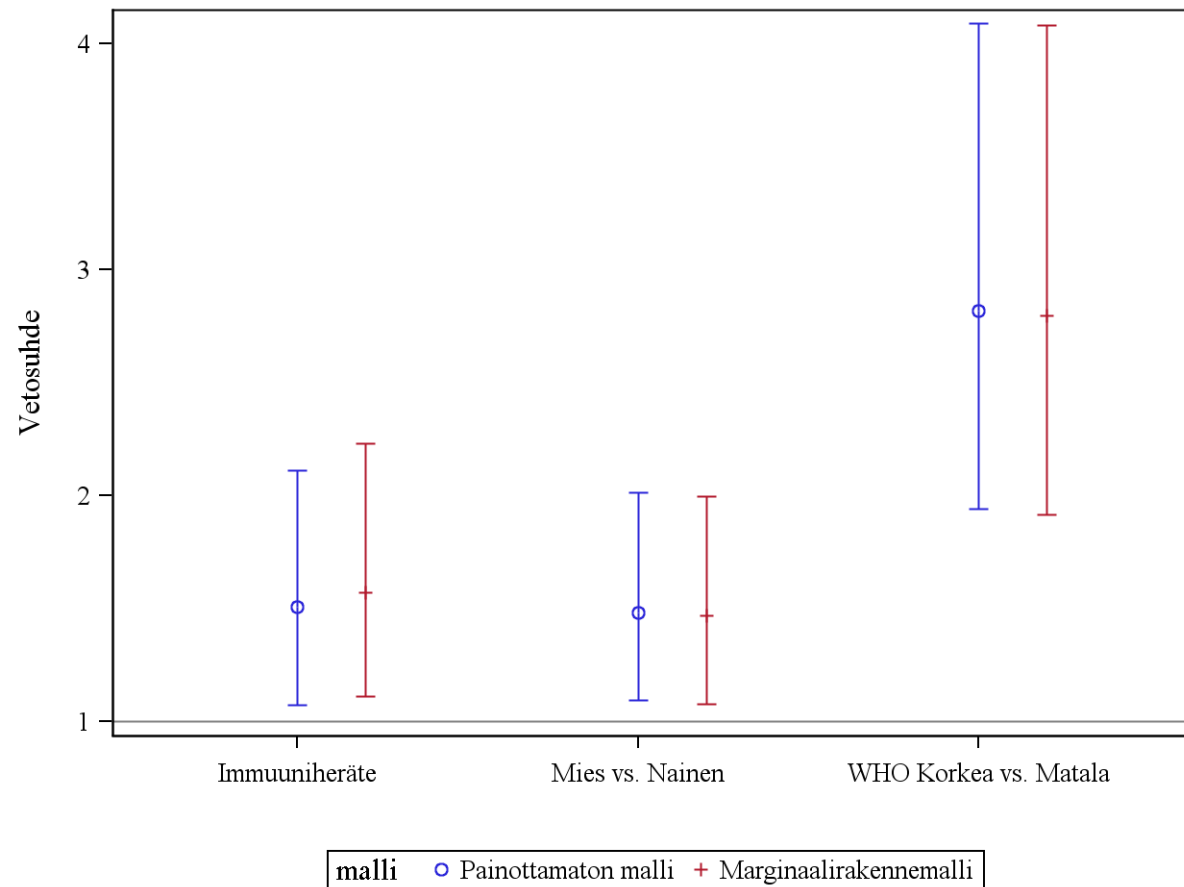
Tulokset ja vertailu



Tulokset ja vertailu

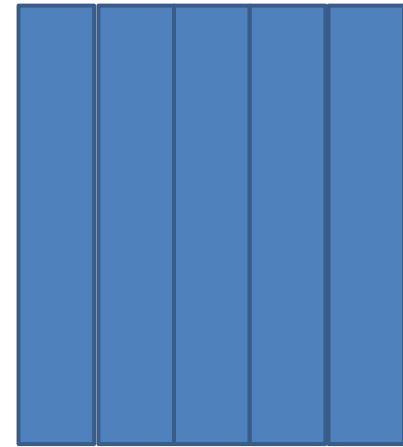
Vertailu	Painottamaton malli		MRM	
	OR	95 % luottamusväli	OR	95 % luottamusväli
Immuuniheräte (Ei vs. Kyllä)	1.506	(1.075, 2.111)	1.573	(1.110, 2.229)
Sukupuoli (mies vs. nainen)	1.483	(1.093, 2.014)	1.467	(1.077, 1.997)
WHO (Korkea vs. Matala)	2.819	(1.942, 4.091)	2.797	(1.917, 4.081)

Tulokset ja vertailu



Miksi näin

- aineistojen havaintojen rakenteet olivat hyvin erilaiset.
 - SOHO-tutkimus: 6516 potilasta ja viisi aikaväliä
 - Virustutkimus: 274 potilasta ja 106 väliä
- SOHO-tutkimuksen Lopullisissa analyyseissä voitiin käyttää monimutkaisempia malleja
- Virusaineistomallissa painojen laskennassa ajauduttiin tyypistämään merkittävästi malleja, jotta estimaattien keskivirheet pysyivät hallinnassa IPT-painoja muodostettaessa





Marginaalirakennemallitus on aikaisempien tutkimusten ja teorian perusteella hyvä työkalu, mutta sen käyttöä ei liene syytä ajatella oletusarvoisena havainnoivien tutkimusten aineistojen analysoinnissa.