

PRO GRADU -TUTKIELMA

Mikko Hietala

**Vakioitavien tekijöiden valinta
lonkkamurtumapotilaiden
hoitokäytäntöjen vertailussa**

TAMPEREEN YLIOPISTO
Matematiikan ja tilastotieteen laitos
Tilastotiede
Huhtikuu 2009

Tampereen yliopisto

Matematiikan ja tilastotieteen laitos

HIETALA, MIKKO: Vakioitavien tekijöiden valinta lonkkamurtumapotilaiden hoitokäytäntöjen vertailussa

Pro gradu -tutkielma, 36 s., 7 liites.

Tilastotiede

Huhtikuu 2009

Tiivistelmä

Lonkkamurtumat ovat merkittävä sairausryhmä, sillä ne heikentävät vanhus-
ten toimintakykyä ja niiden hoidosta aiheutuu paljon kustannuksia terveyden-
huollolle. Lonkkamurtumia havaitaan vuosittain yli 7000 ja määrän on arvioi-
tu kasvavan tulevaisuudessa. Tämän vuoksi olisi tärkeää käyttää taloudelliset
resurssit mahdollisimman tehokkaasti. Tavoitteena voisi olla löytää sellaisia
hoitokäytäntöjä, jotka ovat muita halvempia hoidolle asetettujen tavoitteiden
kuitenkin täyttyessä. Hoitokäytäntöjen vertailussa on kuitenkin huomioitava
erilainen potilasaines. Potilasrakenne saattaa vaihdella alueittain esimerkiksi
ikä- ja sukupuolijakauman osalta. Tutkielmassa etsitään tekijöitä, joiden suh-
teen potilasaines tulisi vakioida, kun halutaan vertailla hoitokäytäntöjä luo-
tettavammin. Itse vakiointia ja hoitokäytäntöjen vertailua ei tässä kuitenkaan
tehdä.

Tutkimusaineisto koostuu 50 vuotta täyttäneistä henkilöistä, jotka ovat
murtaneet lonkkansa ensimmäisen kerran vuosien 1999-2005 välisenä aikana.
Aineisto on Stakesin, Kelan ja yliopistosairaanhoidopiirien yhteistyöhankkeessa
(PERFECT) käytettävä vertailuaineisto. Aineiston tiedot on kerätty eri rekis-
tereistä.

Vakioitavia muuttujia etsitään siten, että muodostetaan logistinen regres-
siomalli. Malliin valitaan sellaisia muuttujia, joilla havaitaan olevan vaikutusta
kuolleisuuteen, joka kuvaa tässä hoidon onnistumista. Havaitaan, että oleellisia
taustamuuttujia ovat ikä, sukupuoli, murtumatyyppi ja muuttuja, joka ilmaisee
onko potilas saanut lyhytaikaista laitoshoidoa murtumaa edeltäneiden kolmen
kuukauden aikana. Lisäksi tarkastellaan potilaiden muita sairauksia. Tärkeim-
miksi liitännäissairauksiksi osoittautuvat syöpä ja sydämen vajaatoiminta. Sai-
rauksia tarkastellaan eri pituisilla ajoilla taaksepäin murtumahetkestä, jotta
saataisiin selville riittävä aika, johon saakka sairauksia kannattaa tarkastella.
Havaitaan, että sairauksien vaikutusten erot eri tarkasteluaikojen välillä ovat
kuitenkin varsin pieniä ja yhtä tiettyä aikaa ei pystytä valitsemaan.

Asiasanat lonkkamurtuma, vakiointi, rekisteri, liitännäissairaus, logistinen
regressio

Sisältö

1 Johdanto	4
1.1 Tutkimuksen tavoitteet	5
1.2 Tutkielman rakenne	6
2 Yleistietoa lonkkamurtumista	7
3 Aineisto	9
4 Vakioitavien tekijöiden valinta	12
4.1 Taustamuuttujat	12
4.2 Liitännäissairaudet	13
5 Logistinen regressio	16
5.1 Määrittely	16
5.2 Kertoimien estimointi	18
5.3 Kertoimien merkitsevyyden testaus	21
5.4 Selittäjien tulkinta	21
5.5 Mallin valinta	23
5.5.1 Devianssi	23
5.5.2 Mallin erottelukyky	24
6 Tulokset	27
6.1 365 päivän kuolleisuus	28
6.2 90 päivän kuolleisuus	31
7 Yhteenveto	34
Lähdeluettelo	36
A Logistisen regression tulosteet	37
B Liitännäissairaudet ja tärkeimpiin sairauksiin liittyvät tilastot	39

1 Johdanto

Vanhusten määrä tulee kasvamaan Suomessa selvästi lähitulevaisuudessa. Samalla kasvavat heidän sairauksistaan terveydenhuollolle aiheutuvat kustannukset. Taloudelliset resurssit ovat kuitenkin rajalliset, joten säästökohteita joudutaan etsimään. Sairaanhoidopiirien hoitokäytäntöjä vertailemalla voitaisiin mahdollisesti löytää hoitokäytäntöjä, jotka ovat muita edullisempia hoidon tason pysyessä kuitenkin riittävän korkeana. Päätöksenteon taustalle tarvitaan kuitenkin luotettavia tilastollisia tuloksia. Jotta sairaanhoidopiirien hoitotulosten vertailu olisi ylipäänsä mahdollista, tarvitaan koko maan kattava luotettava aineisto. Lisäksi on mietittävä, miten saataisiin selville pelkästään hoitokäytäntöjen vaihtelusta aiheutuva vaikutus hoitotuloksiin.

Sosiaali- ja terveystieteiden tutkimus- ja kehittämiskeskus (Stakes), Kansaneläkelaitoksen (Kela) ja yliopistosairaanhoitopiirien yhteishanke PERFECT (PERFormance, Effectiveness and Cost of Treatment episodes eli hoitoketjun toimivuus, vaikuttavuus ja kustannukset) tarkastelee kustannuksiltaan merkittäviä sairausryhmiä (aivohalvaus, lonkkamurtumat, pienipainoiset keskukset, rintasyöpä, skitsofrenia, sydäninfarkti ja tekonivelkirurgia), joiden hoitamisessa erikoissairaanhoidon osuus on suuri. Hankkeen taustalla on Ahvenanmaata lukuun ottamatta koko maan sairaanhoidopiirit kattava aineisto, jonka tiedot on kerätty eri rekistereistä. Rekisteriaineisto mahdollistaa sairaanhoidopiirien vaikuttavuustietojen vertailun, sillä Suomessa terveyteen liittyvät rekisterit ja tilastot ovat varsin monipuolisia ja luotettavia (Sund 2000). Hoitokäytäntöjen vertailua on tarpeellista tehdä paitsi resurssien käytön tehostamisen vuoksi, myös antamaan tietoa hoidon laadusta niin päättäjille kuin kansalaisille (Sund et al. 2008). Tässä tutkielmassa tarkastellaan 50 vuotta täyttäneistä lonkkamurtumapotilaista koostuvaa aineistoa ja tutkitaan, mitä seikkoja täytyy huomioida, jotta hoitokäytäntöjä voidaan vertailla luotettavasti.

Vaikuttavuutta, eli hoidolle asetettujen tavoitteiden toteutumista, voidaan arvioida erilaisilla mittareilla, jotka kuvaavat potilaan terveydentilaa ja sen muutosta. Yleisesti käytetty vaikuttavuusmittari on kuolleisuus. Hoitokäytäntöjä ei voida kuitenkaan suoraan vertailla esimerkiksi kuolleisuuden perusteella, sillä potilasaines saattaa olla erilaista eri alueilla esimerkiksi ikärakenteen osalta. Toisilla alueilla saattaa siis olla vaikeammin hoidettavissa olevia potilaita. Tämä saattaa johtua myös siitä, että eri alueiden hoitoresurssit ovat erilaiset, minkä seurauksena vaikeammin hoidettavissa olevat tapaukset siirretään hoidettavaksi tiettyihin sairaaloihin. Jotta hoitokäytäntöjä pystyttäisiin vertailemaan luotettavammin, täytyy potilasrakenteen erilaisuudesta johtuva vaikutus

eliminoida. Voitaisiin muodostaa ryhmiä, joissa potilaat olisivat taustatiedoilla mahdollisimman samanlaisia. Tällainen järjestely olisi kuitenkin hankalaa varsinkin suuremmassa mittakaavassa, joten avuksi voidaan ottaa tilastolliset menetelmät. Voidaan muodostaa tilastollinen regressiomalli, johon otetaan mukaan sellaisia muuttujia, jotka aiheuttavat ei-toivottua vaihtelua hoitotuloksiin. Tällaista menettelyä kutsutaan mallipohjaiseksi riskivakioinniksi. (Juntunen, Sund, Peltola & Häkkinen 2008.)

1.1 Tutkimuksen tavoitteet

Tässä tutkielmassa etsitään tekijöitä, jotka tulisi vakioida, kun halutaan vertailla lonkkamurtumapotilaiden hoitokäytäntöjä. Erityisen mielenkiinnon kohteena ovat ennen murtumaa potilailla havaitut muut sairaudet. Tavoitteena on löytää sellaiset oheissairaudet, joilla on vaikutusta hoidon onnistumisen kannalta ja jotka siten tulisi huomioida vakioinnissa.

Hoidon onnistumisen mittarina käytetään tässä kuolleisuutta, sillä se on selkeä muuttuja - henkilö on joko kuollut tai elossa tietyn ajan kuluttua lonkkamurtumasta. Oheissairauksien vaikutusta kuolleisuuteen tarkastellaan logistisen regression avulla. Kuolinsyystä ei sinänsä olla kiinnostuneita, aineistossa ei siis erotella johtuuko kuolema lonkkamurtumasta vai ei. Tutkielmassa käytetään vastemuuttujana 90 ja 365 päivän kuolleisuuksia. Kolmen kuukauden kuolleisuuden voisi ajatella ilmaisevan paremmin lonkkamurtumasta johtuvaa kuolleisuutta, mutta lonkkamurtuman vaikutus näkyy myös vuoden kuolleisuudessa. Vuoden kuolleisuuteen vaikuttavat kuitenkin enemmän myös potilaan muut sairaudet.

Koska lonkkamurtumapotilaat ovat yleensä melko iäkkäitä, on heillä tavallisesti myös muita sairauksia, jotka saattavat vaikuttaa kuolleisuuteen. Näitä liitännäissairauksia tarkastellaan lonkkamurtumaa edeltäneeltä ajalta siten, että lyhin tarkastelu-aika on kolme kuukautta ja pisimmillään sairauksia tarkastellaan kymmenen vuotta taaksepäin murtumasta. Tarkoituksena on aluksi löytää sairaudet, joilla on vaikutusta kuolleisuuteen. Sairauksia tarkastellaan taaksepäin eri pituisilla ajoilla, koska halutaan tietää, kuinka pitkälle sairauksia kannattaa tarkastella, eli halutaan löytää lyhin riittävä seuranta-aika oheissairauksille. Aina ei nimittäin saada hoitotietoja niin pitkältä ajalta kuin tässä tutkimuksessa ja aineiston kerääminen esimerkiksi kymmenen vuoden ajalta olisi sitä paitsi turhaa, jos havaitaan lyhyemmänkin ajan riittävän.

Sairauksien tarkasteluajankohdan vaikutusta tutkitaan ensin yksittäisille sairauksille ja sen jälkeen malleille, joissa on mukana useampia sairauksia. Liitännäissairauksien lisäksi selittäjinä käytetään rekistereistä saatavia taustamuuttujia. Tutkielma rajoittuu vakioitavien tekijöiden valintaan, vakiointia ja hoitokäytäntöjen vertailua ei tässä siis tehdä.

1.2 Tutkielman rakenne

Ensimmäisessä luvussa kerrottiin tutkielman aiheesta ja tavoitteista. Toisessa luvussa kerrotaan tarpeellista taustatietoa lonkkamurtumista. Luvussa kolme tutustutaan tarkemmin tutkimusaineistoon. Luvussa neljä kerrotaan vakioinnista ja vakioitavien tekijöiden valinnasta sekä liitännäissairauksien määrittämisestä. Viidennessä luvussa esitellään vakioitavien muuttujien etsimisessä käytettävä tilastollinen menetelmä, logistinen regressio. Luvussa kuusi esitellään saatuja tuloksia. Seitsemännessä luvussa kootaan yhteen päätelmät. Liitteissä on esitelty muodostettujen mallien tuloksia sekä analyyseissa käytetyt oheissairaudet ja oleellisimpiin sairauksiin liittyviä tilastoja.

2 Yleistietoa lonkkamurtumista

Lonkkamurtumat ovat kansantalouden kannalta merkittävä sairausryhmä, sillä murtuma vaatii käytännössä aina leikkauksen ja sen jälkeen usein pitkäaikaista hoitoa ja kuntoutusta terveydenhuollon eri sektoreilla. Lonkkamurtuman jälkeisen vuoden ajalta kertyvät potilaskohtaiset kokonaiskustannukset sairaala- ja laitoshoidossa ovat yli 17 000 euroa, josta erikoissairaanhoidon osuus on noin 23 %. Lonkkamurtuma vaikuttaa suuresti myös potilaan elämänlaatuun, sillä se rajoittaa liikkumista ja saattaa heikentää pysyvästi toimintakykyä. (Sund et al. 2008.)

Lonkkamurtumalla tarkoitetaan reisiluun yläosan murtumaa. Murtumakohdan sijainnin perusteella murtumat jaetaan tavallisesti kolmeen luokkaan, jotka ovat reisiluun kaulan murtumat, trokanteeriset murtumat ja subtrokanteeriset murtumat. Näistä reisiluun kaulan murtuma on yleisin murtumatyyppi. Lähes kaikki lonkkamurtumat tapahtuvat iäkkäille ihmisille. Lonkkamurtuman voidaan ajatella olevan merkki jo alentuneesta toimintakyvystä. Lonkka voi murtua vanhuksilla pienenkin kaatumisen tai putoamisen vuoksi. Lisäksi vanhuksilla kaatumisen riski on suurempi johtuen muista sairauksista ja niiden hoitamiseen tarvittavista lääkkeistä sekä muun muassa näkökyvyn ja tasapainoastin heikkenemisestä (Duodecim 2006). Nuoremmilla lonkkamurtuman syntymekanismikin on erilainen, murtumia aiheuttavat esimerkiksi liikenneonnettomuudet ja korkealta putoamiset. Tämän vuoksi on syytä jättää tarkasteluista pois alle 50-vuotiaat. Yli kaksi kolmasosaa lonkkansa murtaneista on naisia. Tämä johtuu osittain siitä, että naisten keskimääräinen elinikä on korkeampi. (Sund et al. 2008.)

Lonkkamurtuma voidaan hoitaa kiinnittämällä tai proteesin avulla. Leikkausmenetelmän valinta tehdään murtumatyyppin ja potilaan taustatietojen perusteella. Leikkaus tulisi suorittaa mahdollisimman nopeasti murtumisesta, sillä sen on todettu parantavan toipumista ja vähentävän kuolleisuutta (Sund & Liski 2005). Potilaan tulisi kuitenkin olla riittävän hyvässä kunnossa ennen leikkauksen aloittamista, joten liian nopeastikaan ei saa leikata. Lonkkamurtuma ei sinänsä aiheuta kuolemaa, vaan leikkauksen ja toipilasajan aiheuttama yleiskunnon heikkeneminen, joka altistaa potilaan uusille sairauksille. Tärkein kuolleisuuteen vaikuttava tekijä on juuri yleiskunto (Duodecim 2006). Myös leikkauksen jälkeinen kuntoutus tulisi aloittaa mahdollisimman nopeasti ja tehokkaasti. Vaikka tavoitteena on potilaan toimintakyvyn palauttaminen leikkausta edeltävälle tasolle, joutuu suuri osa potilaista pitkäaikaiseen laitoshoitoon murtumaa seuraavan vuoden aikana. Vuonna 2005 uusista lonkkamurtumapotilais-

ta lähes 15 % jäi pitkäaikaispotilaiksi. Noin 20 % lonkkamurtuman kokeneista potilaista on ollut jo murtumahetkellä pitkäaikaisessa laitoshoidossa. Lonkkamurtumapotilaat tarvitsevat keskimäärin lähes 100 hoitopäivää sairaala- ja laitoshoidossa murtumaa seuraavan vuoden aikana. Erikoissairaanhoidon osuus tästä on keskimäärin noin 15 vuorokautta. (Sund et al. 2008.)

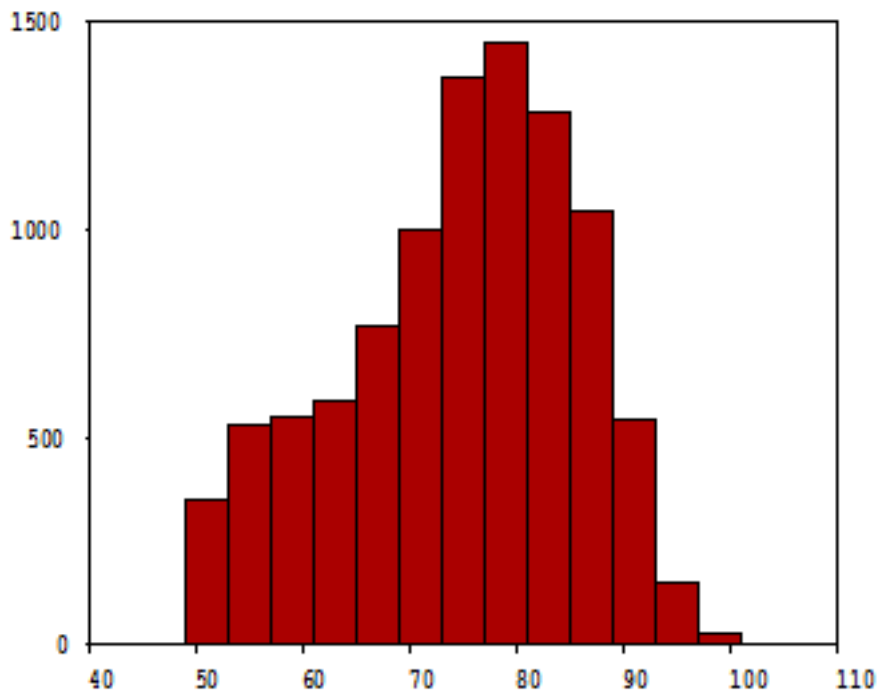
3 Aineisto

Lonkkamurtumapotilaiden hoitokäytäntöjä eri alueiden välillä on vertailtu Suomessa suhteellisen vähän. Suomen kattavat rekisterit antavat tähän kuitenkin hyvät mahdollisuudet. Koska käytännössä kaikki lonkkamurtumat vaativat vuodeosastohoitoa, kirjataan kaikki tapaukset sosiaali- ja terveydenhuollon hoitoilmoitusrekisteriin. Sieltä on poimittu ensin lonkkamurtumadiagnoosin (päätai sivudiagnoosi) saaneet henkilöt. Kunkin potilaan hoitoketju on muodostettu sitten etsimällä tietoja eri rekistereistä henkilötunnuksen perusteella. Hoitoilmoitusrekisteristä saadaan vuodeosastokäyntien hoitotiedot ja lisäksi joitakin taustamuuttujia. Stakesin Sairaaloiden hoitotoiminnan tuottavuus -hankkeen tietokannasta saadaan erikoissairaanhoidon avokäyntien tiedot ja Tilastokeskuksen kuolinsyyrekisteristä potilaiden mahdolliset kuolinpäivämäärät. Lisäksi oheissairauksien määrittelyä varten mukaan on vielä otettu Kelalta saatavat tiedot potilaille myönnettyistä erityiskorvausoikeuksista ja lääkekorvauksista. (Sund et al. 2008.) Hoitoilmoitusjärjestelmästä kertoo enemmän muun muassa Sund (2000) pro gradu -tutkielmassaan.

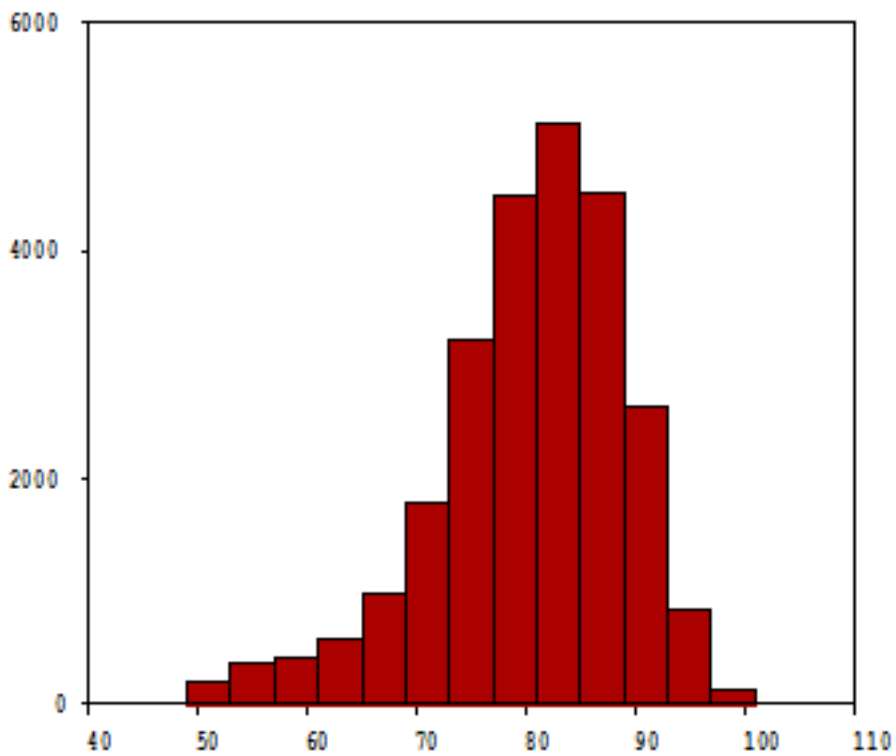
Tässä tutkielmassa käsitellään aineistoa, jossa on mukana murtumahetkellä 50 vuotta täyttäneet ensimmäisen lonkkamurtumadiagnoosin vuosina 1999-2005 saaneet potilaat. Aineisto on PERFECT -yhteistyöhankkeen lonkkamurtuma -osahankkeessa käytettävä vertailuaineisto. Aineistolle tehdyistä rajauksista kertovat tarkemmin Sund et al. (2008).

Rajauksien jälkeen aineiston lopulliseksi kooksi tuli 34 433. Potilaat ovat jakautuneet melko tasaisesti kaikille seitsemälle tarkasteluvuodelle keskiarvon ollessa noin 5000 potilasta per vuosi. Aineiston potilaista 27,5 % on miehiä ja 72,5 % naisia. Potilaat olivat keskimäärin 79 vuoden ikäisiä. Miesten keski-ikä oli 75 vuotta ja naisten 81 vuotta. Kuvioissa 3.1. ja 3.2. on esitettyinä miesten ja naisten ikäjakaumat. Kuvioissa vaaka-akselilla on ikä ja pystyakselilla frekvenssi. Suurin ero kuvioiden muodoissa on noin 50-70-vuotiaiden kohdalla.

Aineiston potilaat on tarkasteluissa jaettu kahteen osaan, omatoimisiin potilaisiin (28 797 kpl) ja pitkäaikaispotilaisiin (5636 kpl). Pitkäaikaispotilaalla tarkoitetaan henkilöä, joka oli saanut pitkäaikaishoidon päätöksen ennen murtumaa tai ollut vähintään 90 päivää yhtäjaksoisesti laitoshoidossa (sairaala, terveyskeskus, vanhainkoti yms.) ennen murtumaa. Loput aineiston potilaista ovat omatoimisia eli kotona asuneita. Pitkäaikaispotilaat on eroteltava omatoimisista potilaista, sillä voidaan olettaa näiden ryhmien potilaiden terveydentilojen eroavan toisistaan. Useimmiten pitkäaikaisesta laitoshoidosta leikkauk-

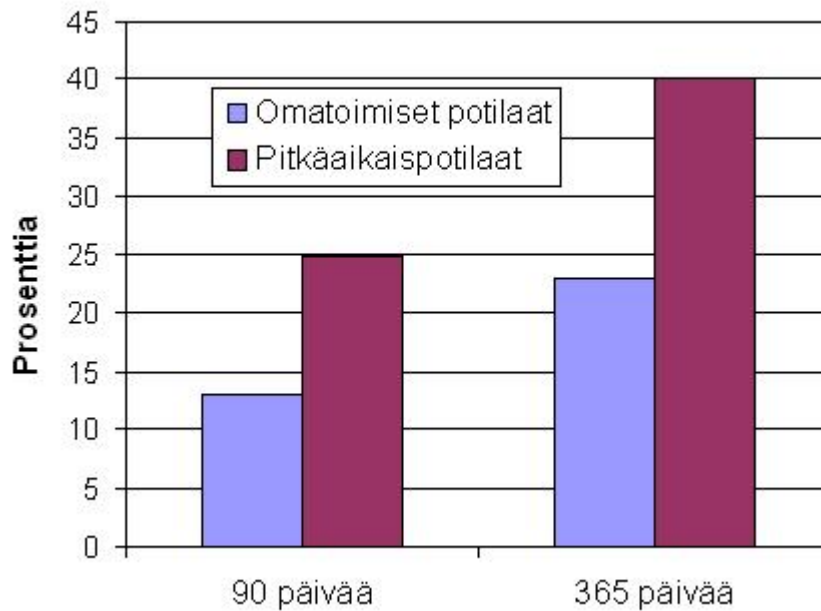


Kuvio 3.1. Miesten ikäjakauma.



Kuvio 3.2. Naisten ikäjakauma.

seen tulevat potilaat ovat huonokuntoisempia kuin kotoa tulevat, mikä lisää kuolleisuusriskiä. Tätä asiaa havainnollistavat erovaisuudet potilaiden tiloissa, kun tarkastellaan leikkauksen jälkeistä aikaa (kuvio 3.3.). Tämän aineiston potilaista kolmen kuukauden kuluttua leikkauksesta omatoimisista potilaista oli kuollut 13 % ja pitkäaikaispotilaista 25 %. Vuoden kohdalla vastaavat luvut olivat omatoimisille 23 % ja pitkäaikaispotilaille 40 %, eli pitkäaikaispotilaiden kuolleisuus oli selvästi suurempaa kummassakin ajankohdassa. Havaittiin, että nämä erot olivat selvästi tilastollisesti merkitseviä, kun testattiin, onko pitkäaikaispotilaiden kuolleisuusosuus suurempi kuin mitä omatoimisten potilaiden joukossa oli havaittu. Omatoimisten ja pitkäaikaispotilaiden vertailu ei kuitenkaan olisi järkevää, sillä laitoshoidossa oleville potilaille ei kirjata lainkaan lääkkeitä eikä sivudiagnooseja merkitä yhtä tarkasti kuin kotona asuville. Tutkielmassa rajoitutaankin tarkastelemaan ainoastaan omatoimisia potilaita.



Kuvio 3.3. Omatoimisten ja pitkäaikaispotilaiden kuolleisuusosuudet 90 ja 365 päivän kuluttua leikkauksesta.

4 Vakioitavien tekijöiden valinta

Hoitotuloksiin vaikuttaa yleisesti ottaen kolme tekijää: potilaskohtaiset tekijät, tuottajakohtaiset tekijät (hoitokäytännöt) sekä satunnaisvaihtelu. Nyt olemme kiinnostuneita hoitokäytännöistä aiheutuvasta vaikutuksesta. Potilaskohtainen vaihtelu, joka voidaan ajatella sekoittavaksi tekijäksi hoitotulosten tarkastelun kannalta, ei ole niin kiinnostavaa ja tästä vaikutuksesta halutaankin eroon. (Juntunen et al. 2008.)

Kun halutaan vertailla hoitotuloksia, potilasjoukkojen tulisi olla mahdollisimman homogeenisia terveydentilojen ja sosiodemografisten tekijöiden suhteen (Holman, Preen, Baynham, Finn, & Semmens 2005). Jos näin ei ole, tulokset vääristyvät. Tämä ongelma voidaan ratkaista vakioimalla potilasaines tilastollisin menetelmin. Vakiointi voidaan tehdä muodostamalla tilastollinen regressiomalli, johon valitaan sellaisia muuttujia, jotka aiheuttavat ei-toivottua vaihtelua. Tässä tutkielmassa ei kuitenkaan tehdä itse vakiointia, vaan tarkoituksena on ainoastaan etsiä tekijöitä, joiden suhteen vakiointi tulisi suorittaa.

Juntunen et al. (2008) havaitsivat, että jo yksinkertainen riskivakiointimalli aiheutti merkittäviä eroja vaikuttavuusluvuista tehtyihin päätelmiin, kun tarkastelun kohteena olivat aivoinfarktiin sairastuneet potilaat. Vakiointi on siis tarpeellista ja vakioitavien muuttujien valinta on siten tehtävä huolellisesti. Jos mallista puuttuu olennaisia muuttujia, saatetaan osa potilasaineksen vaihtelusta tulkita virheellisesti hoitokäytäntöjen eroina. Ei ole myöskään hyvä, että mallissa on liikaa selittäjiä, vaan mukana tulisi olla ainoastaan olennaiset muuttujat.

Vakioitavia tekijöitä etsitään siten, että muodostetaan logistinen regressiomalli. Malliin valitaan sellaisia potilaskohtaisia muuttujia, joilla on vaikutusta kuolleisuuteen, joka kuvastaa tässä hoitotuloksia. Vakioimalla näiden muuttujien vaikutus pystytään hoitokäytäntöjä vertailemaan luotettavammin. Seuraavissa alaluvuissa esitellään tekijöitä, joiden vaikutusta kuolleisuuteen tutkitaan luvussa kuusi.

4.1 Taustamuuttujat

Taustamuuttujiksi kutsutaan tässä muita kuin oheissairauksiin liittyviä muuttujia, joiden valintaan pääasiassa keskitytään. Yhteistä näille taustamuuttujille on se, että ne ovat määräytyneet jo ennen hoidon aloittamista, toisin sanoen muuttujat eivät ole osa hoitokäytäntöjä. Valintaa rajoittaa kuitenkin se, että näiden muuttujien on löydettävä rekistereistä. Tämän vuoksi rekisteriaineistoa

käytettäessä vakiointia ei välttämättä pystytä tekemään kaikkien haluttujen muuttujien suhteen.

Tärkeitä taustamuuttujia ovat ainakin ikä ja sukupuoli. Myös se, onko potilas ollut lyhytaikaisessa laitoshoidossa murtumaa edeltäneiden kolmen kuukauden aikana (1-89 päivää), vaikuttaa oletettavasti kuolleisuuteen. Murtumatyyppien vaikutusta on myös syytä tutkia. Suurin mielenkiinto kohdistuu kuitenkin potilaiden liitännäissairauksiin, joista kerrotaan tarkemmin seuraavassa alaluvussa.

4.2 Liitännäissairaudet

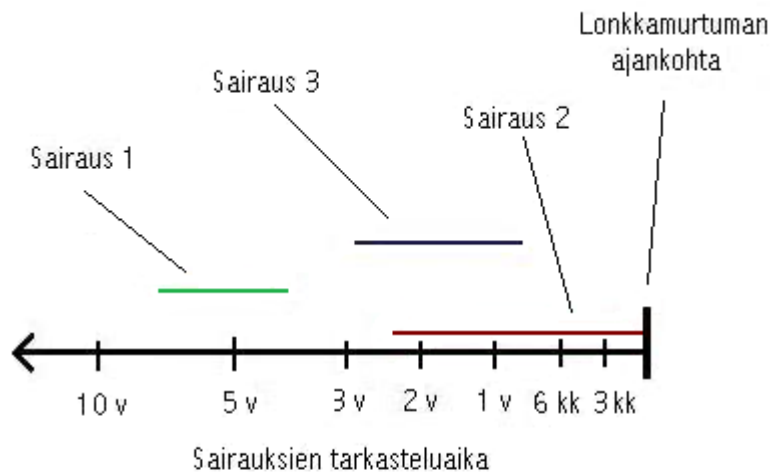
Lonkkamurtumapotilailla on tavallisesti myös muita sairauksia samanaikaisesti tai on ollut muita sairauksia ennen murtumaa. Näitä liitännäissairauksia kutsutaan usein myös komorbiditeeteiksi. Komorbiditeetit vaikuttavat usein siihen, pystytäänkö potilas kuntouttamaan ennalleen tavoitteen mukaisesti. Huono yleiskunto saattaa hidastaa toipumista, josta voi aiheutua potilaan joutuminen pitkäaikaiseen laitoshoitoon tai jopa kuolema. On selvää, että mitä paremmassa kunnossa potilas on, sitä paremmin hän keskimäärin toipuu leikkauksesta.

Potilaiden terveydentilat ja edeltävät sairaudet saattavat olla erilaisia eri alueilla. Nämä erot on syytä huomioida, kun halutaan vertailla alueiden hoitotuloksia keskenään. Tavoitteena on löytää kuolleisuuteen merkittävästi vaikuttavat oheissairaudet. Sairauksia tarkastellaan eri ajan pituuksilla taaksepäin murtumasta, jotta saadaan selville riittävä tarkasteluaika. Tarkasteltavat ajan pituudet ovat 3 kuukautta, 6 kuukautta, 1 vuosi, 2 vuotta, 3 vuotta, 5 vuotta ja 10 vuotta. Kuukauden pituudeksi on määritelty 30 päivää ja vuoden pituudeksi 365 päivää. Myös liitännäissairauksien ja iän sekä sukupuolen välisiä yhdysvaikutuksia eli interaktioita on syytä tarkastella, sillä voi olla että jonkun sairauden vaikutus kuolleisuuteen on esimerkiksi suurempi miehillä kuin naisilla tai suurempi nuoremmilla verrattuna vanhempiin.

Komorbiditeettien valinta ja luokittelu perustuu Charlsonin, Elixhauserin ja Jiangin luokitukseen (Holman et al. 2005; Quan et al. 2005; Jiang et al. 2005). Komorbiditeettien määrittely on aluksi tehty hoitoilmoitusrekisterin ICD-10 (nykyinen luokitus, käytössä vuodesta 1996) ja ICD-9 -tautiluokituskoodein perusteella. Kuhunkin sairausryhmään liittyy useita ICD -koodeja. Sairauksien identifioimiseksi on lisäksi määritelty mitkä erityiskorvausoikeudet ja lääkeostot liittyvät mihinkin sairausryhmään. Erityiskorvausoikeus myönnetään kroonisiin sairauksiin, eli sairauksiin, joista ei käytännössä parannuta. Tämän vuoksi korvausoikeudelle ei ole yleensä määritelty päättymisajankohtaa. Sen sijaan esimerkiksi syövästä voi parantua täysin, jolloin sen esiintyminen vaikkapa viisi vuotta sitten ei enää välttämättä nosta kuolleisuusriskiä.

Tarkastellaan kuvion avulla yksinkertaista tilannetta, joka voisi kuvata yksittäisen potilaan sairaushistoriaa. Kuviossa 4.1. potilaalla on ollut kolme eri sairautta, joiden vaikutusta on havainnollistettu viivojen avulla. Viivojen va-

sen pää kuvaa ensimmäistä ja oikea pää viimeistä tapahtumaa. Vaaka-akselilla on oheissairauksien tarkasteluajat. Sairaus 1 havaitaan, kun tarkastelu aika on vähintään 5 vuotta, sillä viimeinen hoitokerta on ollut noin neljä vuotta ennen lonkkamurtumaa. Kyseiseen sairauteen ei ole kuitenkaan tarvittu hoitoa sen jälkeen. Sairauteen 2 liittyy erityiskorvausoikeus, joka on edelleen voimassa lonkkamurtuman tapahtuessa. Tällaiset krooniset sairaudet havaitaan siis jo lyhimmillä (kolmen kuukauden) seuranta-ajalla. Sairauteen 3 liittyviä lääkkeitä henkilö on tehnyt viimeisen kerran noin 9 kuukautta ennen murtumaa. Tämä sairaus havaitaan siis ensimmäisen kerran vuoden tarkasteluajalla. Liitännäissairauksiin voi tietysti liittyä esimerkiksi sekä hoitoilmoituksia että lääkkeitä. Tällöin sairauden esiintymisajankohta määritellään viimeisen tapahtuman perusteella riippumatta siitä, millä tavalla se on havaittu. Kaikkiin komorbiditeetteihin voi liittyä hoitoilmoituksia, mutta erityiskorvausoikeus koskee ainoastaan kahtatoista ja lääkekorvaukset kymmentä komorbiditeettiä.



Sairaus 1: ensimmäisen ja viimeisen hoitokerran välinen aika

Sairaus 2: erityiskorvausoikeuden alkamis- ja päättymiskohta

Sairaus 3: ensimmäinen ja viimeinen lääkeostokerta

Kuvio 4.1. Kuvitteellisen potilaan sairaushistoria.

Tässä tutkielmassa tarkastellaan viimeisiä tapahtumia (hoitoilmoitukset, erityiskorvausoikeudet ja lääkeostot) ennen indeksihoitojakson alkua. Indeksihoidojaksolla tarkoitetaan tässä ensimmäistä lonkkamurtuman hoitojaksoa. Hoitoilmoitukset sisältävät hoitotiedot sekä perusterveydenhuollon (tiedot ainoastaan vuodeosastohoidosta) että erikoissairaanhoidon osalta. Oheissairauksien havaitsemiseen on käytetty sekä pää- että sivudiagnooseja. Jos viimeinen hoitomerkinä on tarkasteltavalla ajanjaksolla, on potilaan katsottu saaneen

hoitoa kyseiseen sairauteen. Erityiskorvausosoikeuksista tarkastellaan korvausosoikeuden päättymisajankohtaa tai voimassaoloa. Jos korvausosoikeus on edelleen voimassa, katsotaan potilaan sairastavan edelleen kyseistä sairautta. Tämän vuoksi erityiskorvausosoikeuksien avulla havaittujen tapausten määrä ei juurikaan lisääny, kun tarkasteluaikaa pidennetään. Lääkeostoista tarkastellaan kuhunkin sairauteen liittyvää viimeistä ostokertaa.

Liitännäissairaudet on määritelty hoitoilmoitusrekisterin, erityiskorvattavuuksien ja lääkeostojen perusteella siten, että jos potilaalla on tiettyyn sairauteen liittyvä merkintä vähintään yhdessä näistä rekistereistä, katsotaan hänen sairastaneen kyseisen sairauden. Jos siis tarkastellaan sairauksia esimerkiksi vuoden ajalta, täytyy jonkun seuraavista kriteereistä toteutua: 1) viimeisin hoitoilmoitus (pää- tai sivudiagnoosi) tarkasteltavasta sairaudesta on korkeintaan vuoden takaa ennen lonkkamurtuman indeksijakson alkua, 2) erityiskorvausosoikeus on edelleen voimassa tai ollut voimassa vielä vuosi sitten tai 3) viimeisestä kyseiseen sairauteen liittyvästä lääkeostokerrasta on korkeintaan vuosi. Indeksihoidojakson alkamisen jälkeisiä diagnooseja, erityiskorvausosoikeuksien alkamisia ja lääkeostoja ei ole huomioitu. Mukaan on valittu 32 liitännäissairautta (ks. liite B), joiden vaikutusta kuolleisuuteen tarkastellaan logistisen regression avulla. Sen teoria esitellään seuraavassa luvussa.

5 Logistinen regressio

Tämän luvun lähteenä (lukuun ottamatta alalukua 5.5.2) käytetään Hosmerin ja Lemeshow'n kirjaa *Applied logistic regression* (1989) ja pääasiassa sen lukuja 1 ja 3.

Logistinen regressio(analyysi) on yleisesti käytetty menetelmä tutkittaessa esimerkiksi terveystieteellisiä aineistoja, joissa selitettävä muuttuja on usein kaksiluokkainen (dikotominen) eli voi saada ainoastaan kaksi arvoa (tavallisesti arvot 0 ja 1). Logistinen regressio kuuluu yleistettyjen lineaaristen mallien perheeseen, kuten myös tavallinen lineaarinen regressioanalyysi. Voisimme käyttää myös lineaarista regressiota selittämään dikotomista vastemuuttujaa, mutta se ei kuitenkaan yleensä sovellu yhtä hyvin tähän tarkoitukseen. Linearisessa regressiossa nimittäin selitettävä muuttuja oletetaan jatkuvaksi. Toinen tärkeä ero on siinä, että lineaarinen regressio pyrkii selittämään määriä, logistinen regressio puolestaan todennäköisyyksiä. Muilta osin lineaarisen regression menetelmät ja periaatteet soveltuvat varsin hyvin käytettäväksi logistisessa regressiossa. Tarkastellaan jatkossa tarkemmin sekä näiden menetelmien eroja että yhtäläisyyksiä.

Logistisessa regressiossa vasteena on usein jokin ilmiö, joka tapahtuu (arvo 1) tai ei tapahdu (arvo 0). Logistinen regressio pyrkii ennustamaan näiden arvojen todennäköisyyksiä kullekin havainnolle. Todennäköisyysestimaatit määntyvät regressiomallin sisältämien selittäjien kertoimien estimaattien ja havaintoarvojen perusteella. Selittäjät, joita kutsutaan usein kovariaateiksi, ovat myös usein dikotomisissa, mutta voivat olla myös useampiluokkaisia tai jatkuvia. Selitettävä muuttuja voi myös saada useampia arvoja. Tällöin käytetään multinomiaalia logistista regressiota. Tässä tutkielmassa käytetään kuitenkin tavallista binääristä logistista regressiota, sillä vasteena on kaksiluokkainen muuttuja kuolleisuus. Vastemuuttuja saa arvon 1, jos potilas on kuollut ja arvon 0, jos potilas on elossa tarkasteltavana ajankohtana.

5.1 Määrittely

Logistista regressiota voidaan lähestyä tavallisen regressioanalyysin kautta. Lineaarinen regressio oletetaan pääpiirteittäin tutuksi lukijalle. Tarkastellaan jatkossa yksinkertaista tilannetta, jossa meillä on ainoastaan yksi selittäjä. Määritelmät ovat kuitenkin helposti yleistettävissä usean selittäjän tilanteeseen.

Määritellään aluksi lineaarinen regressiomalli siten, että

$$Y = \beta_0 + \beta_1 x + \epsilon,$$

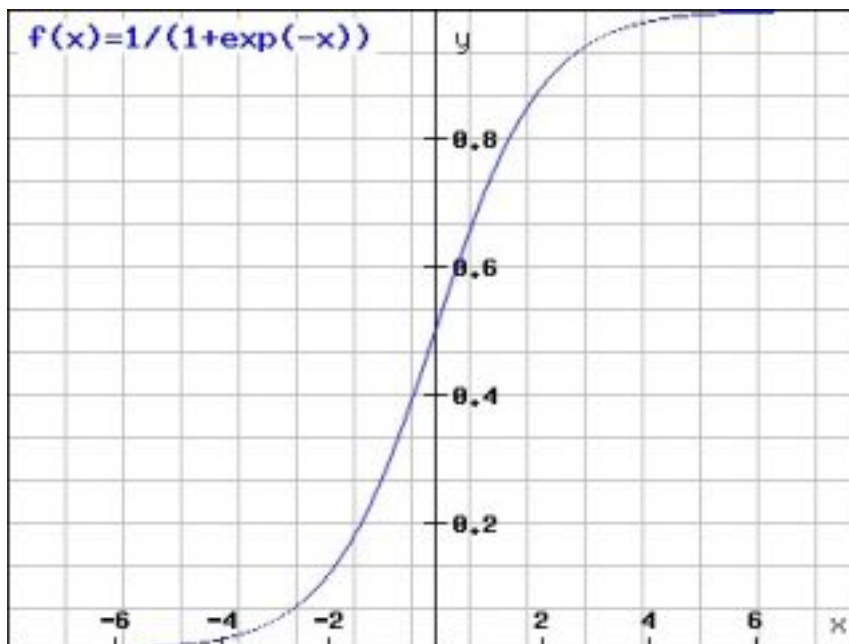
missä Y on vastemuuttuja, x tarkoittaa selittäjän X arvoa ja virhetermi $\epsilon \sim N(0, \sigma^2)$. Vasteen keskimääräinen arvo eli ehdollinen odotusarvo voidaan esittää muodossa $E(Y|x)$. Lineaarisen regressiomallin tapauksessa

$$E(Y|x) = \beta_0 + \beta_1 x.$$

Lauseke voi saada mitä tahansa arvoja, kun x vaihtelee välillä $(-\infty, \infty)$. Logistisen regression tapauksessa ennustetaan todennäköisyyksiä, joten käytettävän mallin tulisi kuitenkin antaa arvoja väliltä $[0, 1]$. Selitettävän muuttujan ollessa dikotominen ei lineaarinen regressiomalli myöskään kuvaa hyvin vasteen ja selittäjien välistä suhdetta. Logistisessa regressiossa odotusarvosta $E(Y|x)$ käytetään merkintää $\pi(x)$ ja funktiona $\pi(x)$ käytetään tässä niin sanottua logistista funktiota, jolloin

$$(5.1) \quad \pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}.$$

Logistisen funktion valintaa voidaan perustella muun muassa siten, että se on matemaattisesti ajateltuna joustava ja helppokäyttöinen ja sitä pidetään myös järkevästi tulkittavana. Voitaisiin käyttää myös esimerkiksi probit-funktiota, joka tuottaa yleensä melko samanlaisia tuloksia kuin tässä käytetty logit-funktio. Logistinen funktio $\pi(x)$ saa nyt arvoja väliltä $[0, 1]$. Suurilla negatiivisilla x :n arvoilla $\pi(x)$ saa lähellä nollia olevia arvoja ja suurilla positiivisilla x :n arvoilla se saa lähellä ykköstä olevia arvoja. Vasteen ja selittäjien välinen suhde graafisesti esitettynä on nyt s-kirjaimen muotoinen (kuvio 5.1.).



Kuvio 5.1. Logistinen funktio.

Myöskään vasteen ja virhetermin jakaumille asetetut oletukset eivät välttämättä ole voimassa, kun vastemuuttuja on kaksiluokkainen. Linearisessa

regressiomallissa yksittäinen Y :n havaintoarvo voidaan esittää kaavan $y = E(Y|x) + \epsilon$ mukaan. Yleensä oletamme, että virhetermi ϵ noudattaa normaali-jakaumaa keskiarvon ollessa 0 ja varianssina jokin vakio. Tällöin myös y noudattaa normaalijakaumaa keskiarvon ollessa $E(Y|x)$ ja varianssina vakio.

Logistisen regressiomallin tapauksessa vasteen arvot voidaan esittää muodossa $y = \pi(x) + \epsilon$. Virhetermi ϵ ei ole nyt normaalisti jakautunut, sillä se voi saada ainoastaan kaksi arvoa: jos $y = 1$, niin $\epsilon = 1 - \pi(x)$ ja jos $y = 0$, niin $\epsilon = -\pi(x)$. Näille tapahtumille voidaan määritellä myös todennäköisyydet siten, että $y = 1$ todennäköisyydellä $\pi(x)$ ja $y = 0$ todennäköisyydellä $1 - \pi(x)$. Virhetermi noudattaa nyt binomijakaumaa keskiarvon ollessa 0 ja varianssin $\pi(x)[1 - \pi(x)]$. Täten myös vastemuuttujan arvojen jakauma on binomijakautunut odotusarvon ollessa nyt $\pi(x)$ ja varianssin sama kuin edellä, $\pi(x)[1 - \pi(x)]$. Varianssi ei siis ole vakio, kuten lineaarisen regression puolella oletetaan.

Logistisessa regressiossa käytetään linkkifunktiona (funktio, jonka avulla voidaan esittää vasteen odotusarvon ja mallin lineaarisen osan yhteys) niin sanottua logit-muunnosta. Muunnos voidaan esittää kaavan (5.1) ja luonnollisen logaritmin määritelmän perusteella siten, että

$$\begin{aligned} g(x) &= \ln \left[\frac{\pi(x)}{1 - \pi(x)} \right] = \ln(\pi(x)) - \ln(1 - \pi(x)) \\ &= \ln \left[\frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}} \right] - \ln \left[1 - \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}} \right] \\ &= \ln(1) - \ln(1 + e^{-(\beta_0 + \beta_1 x)}) - \ln \left[\frac{e^{-(\beta_0 + \beta_1 x)}}{1 + e^{-(\beta_0 + \beta_1 x)}} \right] \\ &= -\ln(1 + e^{-(\beta_0 + \beta_1 x)}) - \ln(e^{-(\beta_0 + \beta_1 x)}) + \ln(1 + e^{-(\beta_0 + \beta_1 x)}) \\ &= -\ln(e^{-(\beta_0 + \beta_1 x)}) = \beta_0 + \beta_1 x. \end{aligned}$$

Tämän perusteella voidaan määritellä yksinkertainen logistinen regressiomalli

$$\ln \left[\frac{\pi(x)}{1 - \pi(x)} \right] = \beta_0 + \beta_1 x.$$

Tämä malli voidaan ajatella tavalliseksi regressiomalliksi, jossa yhtälön vasen puoli toimii nyt selitettävänä muuttujana. Logit-muunnoksen hyödyllisyys on siinä, että funktio $g(x)$ toteuttaa monia toivottavia ominaisuuksia, joita asetetaan lineaariselle regressiomallille. Ensinnäkin $g(x)$ on lineaarinen parametriensa suhteen, se on jatkuva ja voi myös saada arvoja väliltä $(-\infty, \infty)$ riippuen x :n arvoista.

5.2 Kertoimien estimointi

Oletetaan, että meillä on n :n riippumattoman havainnon otos havaintoparisista (x_i, y_i) , $i = 1, 2, \dots, n$, missä y_i on dikotomisen vastemuuttujan arvo (0 tai 1) ja x_i riippumattoman selittäjän arvo i . havainnolle. Tavoitteena on estimoida regressiokertoimet β_0 ja β_1 yhtälöstä (5.1). Linearisessa regressiossa

kertoimet estimoidaan yleensä pienimmän neliösumman menetelmällä. Tätä menetelmää ei voida kuitenkaan käyttää silloin, kun selitettävä muuttuja on dikotominen. Logistisessa regressiossa estimointiin käytetään yleensä suurimman uskottavuuden menetelmää, jonka pääperiaatteet ja tarvittavat funktiot esitellään seuraavaksi.

Muodostetaan aluksi yhteisjakauman todennäköisyysfunktio havaintoparille (x_i, y_i) :

$$\xi(x_i) = \pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i}.$$

Lauseke muodostuu edellä esitetyistä todennäköisyyksistä $P(y = 1) = \pi(x)$ ja $P(y = 0) = 1 - \pi(x)$, missä $\pi(x)$ on määritelty yhtälössä (5.1).

Muodostetaan seuraavaksi niin sanottu uskottavuusfunktio, joka ilmaisee tuntemattomien parametrien funktiona todennäköisyyden, että saadaan havaittu data. Uskottavuusfunktio saadaan termien $\xi(x_i)$ tulona yli otoskoon n , eli

$$(5.2) \quad l(\boldsymbol{\beta}) = \prod_{i=1}^n \xi(x_i),$$

missä vektori $\boldsymbol{\beta}$ sisältää parametrit β_0 ja β_1 .

Suurimman uskottavuuden menetelmän periaatteena on valita parametrivektorin $\boldsymbol{\beta}$ estimaatiksi arvo, joka maksimoi uskottavuusfunktion. Matemaattisesti on kuitenkin helpompaa käyttää yhtälön (5.2) logaritmia. Logaritmoitu uskottavuusfunktio on tällöin

$$(5.3) \quad L(\boldsymbol{\beta}) = \ln[l(\boldsymbol{\beta})] = \sum_{i=1}^n \{y_i \ln[\pi(x_i)] + (1 - y_i) \ln[1 - \pi(x_i)]\}.$$

Löytääksemme arvot, jotka maksimoivat yhtälön (5.3), derivoidaan kyseinen yhtälön parametrien β_0 ja β_1 suhteen. Saadaan seuraavat yhtälökettjut:

$$\begin{aligned} \frac{\partial L(\boldsymbol{\beta})}{\partial \beta_0} &= \sum_{i=1}^n \left[y_i \frac{\partial}{\partial \beta_0} \ln[\pi(x_i)] + (1 - y_i) \frac{\partial}{\partial \beta_0} \ln[1 - \pi(x_i)] \right] \\ &= \sum_{i=1}^n \left[y_i \frac{\partial}{\partial \beta_0} \ln \left[\frac{1}{1 + e^{-\beta_0 - \beta_1 x_i}} \right] + (1 - y_i) \frac{\partial}{\partial \beta_0} \ln \left[1 - \frac{1}{1 + e^{-\beta_0 - \beta_1 x_i}} \right] \right] \\ &= \sum_{i=1}^n \left[y_i \left[\frac{e^{-\beta_0 - \beta_1 x_i}}{1 + e^{-\beta_0 - \beta_1 x_i}} \right] + (1 - y_i) \left[-1 + \frac{e^{-\beta_0 - \beta_1 x_i}}{1 + e^{-\beta_0 - \beta_1 x_i}} \right] \right] \\ &= \sum_{i=1}^n \left[y_i - \frac{1}{1 + e^{-\beta_0 - \beta_1 x_i}} \right] = \sum_{i=1}^n \left[y_i - \pi(x_i) \right] \end{aligned}$$

ja

$$\begin{aligned}
\frac{\partial l(\boldsymbol{\beta})}{\partial \beta_1} &= \sum_{i=1}^n \left[y_i \frac{\partial}{\partial \beta_1} \ln[\pi(x_i)] + (1 - y_i) \frac{\partial}{\partial \beta_1} \ln[1 - \pi(x_i)] \right] \\
&= \sum_{i=1}^n \left[y_i \frac{\partial}{\partial \beta_1} \ln \left[\frac{1}{1 + e^{-\beta_0 - \beta_1 x_i}} \right] + (1 - y_i) \frac{\partial}{\partial \beta_1} \ln \left[1 - \frac{1}{1 + e^{-\beta_0 - \beta_1 x_i}} \right] \right] \\
&= \sum_{i=1}^n \left[y_i x_i \left[\frac{e^{-\beta_0 - \beta_1 x_i}}{1 + e^{-\beta_0 - \beta_1 x_i}} \right] + (1 - y_i) x_i \left[-1 + \frac{e^{-\beta_0 - \beta_1 x_i}}{1 + e^{-\beta_0 - \beta_1 x_i}} \right] \right] \\
&= \sum_{i=1}^n x_i \left[y_i - \frac{1}{1 + e^{-\beta_0 - \beta_1 x_i}} \right] = \sum_{i=1}^n x_i \left[y_i - \pi(x_i) \right].
\end{aligned}$$

Asettamalla nämä derivaatat nolliksi saadaan seuraavat niin sanotut uskottavuusyhtälöt:

$$\sum_{i=1}^n \left[y_i - \pi(x_i) \right] = 0$$

ja

$$\sum_{i=1}^n x_i \left[y_i - \pi(x_i) \right] = 0.$$

Nämä yhtälöt eivät ole lineaarisia parametrien $\boldsymbol{\beta}$ suhteen, joten ne eivät ole suoraviivaisesti ratkaistavissa. Estimaattien saamiseksi näistä yhtälöistä on käytettävä iteratiivisia menetelmiä, joita ei tässä kuitenkaan esitellä, sillä estimaatit saadaan laskettua helposti tietokoneohjelmilla. Uskottavuusyhtälöiden ratkaisuna saadaan suurimman uskottavuuden estimaatit, joiden vektoria merkitään symbolilla $\hat{\boldsymbol{\beta}}$. Näiden estimaattien avulla saadaan estimaatti todennäköisyydelle $\pi(x)$, jolloin

$$\hat{\pi}(x) = \frac{1}{1 + e^{-\hat{g}(x)}},$$

missä $\hat{g}(x) = \hat{\beta}_0 + \hat{\beta}_1 x$.

Kertoimille saadaan laskettua myös estimoidut keskihajonnat, joita tarvitaan luottamusvälejä muodostettaessa sekä kertoimien merkitsevyyttä testattaessa. Estimaatit saadaan logaritmoidun uskottavuusfunktion toisen kertaluvun osittaisderivaattojen avulla.

Kertoimille saadaan määriteltyä $100 \times (1 - \alpha)\%$ luottamusväli siten, että alaraja

$$LV_a = \hat{\beta}_i - z_{1-\alpha/2} * \hat{SE}(\hat{\beta}_i)$$

ja yläraja

$$LV_y = \hat{\beta}_i + z_{1-\alpha/2} * \hat{SE}(\hat{\beta}_i).$$

Lausekkeissa $\hat{SE}(\hat{\beta}_i)$ on i. kertoimen estimaatin estimoitu keskihajonta ja z normaali jakauman taulukkoarvo merkitsevyytätasolla α .

5.3 Kertoimien merkitsevyyden testaus

Muuttujien tarpeellisuutta regressiomallissa voidaan arvioida testaamalla saatujen kerroinestimaattien merkitsevyyttä. Testaus on samantapaista kuin lineaarisessa regressiossa, eli asetetaan ensin hypoteesit ja valitaan jokin merkitsevyydestaso, jolla nollahypoteesi hylätään. Seuraavaksi lasketaan testisuureen arvo ja sitä vastaava p-arvo. Jos saatu p-arvo on pienempi kuin valittu merkitsevyydestaso, nollahypoteesi hylätään ja voidaan sanoa, että tarkasteltava selittäjä selittää tilastollisesti merkitsevän osan vasteesta. Muussa tapauksessa nollahypoteesi jää voimaan. Logistisessa regressiossa kertoimien testaamiseen voidaan käyttää monia erilaisia testejä, joista tässä käytämme niin sanottua Waldin χ^2 -testiä.

Waldin testisuure saadaan jakamalla kertoimen estimaatti sen keskihajonnan estimaatilla ja korottamalla tämä suhde toiseen potenssiin. Testisuure noudattaa likimain χ^2 -jakaumaa vapausastein 1, eli

$$W = \left(\frac{\hat{\beta}_1}{\widehat{SE}(\hat{\beta}_1)} \right)^2 \sim \chi^2(1).$$

Nollahypoteesina testissä on $H_0 : \hat{\beta}_1 = 0$ ja vaihtoehtoisena hypoteesina $H_1 : \hat{\beta}_1 \neq 0$. H_0 hylätään, jos p-arvo on pienempi kuin valittu merkitsevyydestaso α , joka on tavallisesti 0,05 eli 5 %. Tällöin hyväksytään H_1 ja voidaan sanoa, että kerroin eroaa tilastollisesti merkitsevästi nolasta. Waldin testistä voidaan käyttää myös muotoa, jossa kertoimen ja sen keskihajonnan suhdetta ei koroteta toiseen potenssiin. Tällöin saadaan tavanomainen lineaarisessa regressiossa käytettävä testisuure, joka noudattaa normaalijakaumaa.

5.4 Selittäjien tulkinta

Oletetaan, että sekä selitettävä että selittävä muuttuja (Y ja X) ovat dikotomisia ja voivat saada arvon 0 tai 1. Näin meillä on yhteensä neljä mahdollista arvoa todennäköisyyksille $\pi(x)$ ja $1 - \pi(x)$ riippuen x:n ja y:n arvoista. Nämä todennäköisyydet on määritelty taulukossa 5.1.

Taulukko 5.1. Logistisen regressiomallin todennäköisyysarvot, kun muuttujat Y ja X dikotomisia.

Arvot	$x = 1$	$x = 0$
$y = 1$	$\pi(1) = \frac{1}{1+e^{-(\beta_0+\beta_1)}}$	$\pi(0) = \frac{1}{1+e^{-\beta_0}}$
$y = 0$	$1 - \pi(1) = \frac{e^{-(\beta_0+\beta_1)}}{1+e^{-(\beta_0+\beta_1)}}$	$1 - \pi(0) = \frac{e^{-\beta_0}}{1+e^{-\beta_0}}$

Todennäköisyyksien perusteella tapahtumalle voidaan määritellä sen veto (odds). Se saadaan jakamalla tapahtuman toteutumisen todennäköisyys todennäköisyydellä, että tapahtuma ei toteudu, eli

$$odds = \frac{\pi(x)}{1 - \pi(x)}.$$

Kun tästä lausekkeesta otetaan luonnollinen logaritmi, päädytään itse asiaansa logit-muunnokseen

$$g(x) = \ln \left[\frac{\pi(x)}{1 - \pi(x)} \right].$$

Muuttujan X arvojen vaikutusten suhdetta vasteeseen voidaan tarkastella niin sanotun vetosuhteen (odds ratio, OR) avulla. Se saadaan jakamalla arvojen $x=1$ ja $x=0$ vedot keskenään, eli

$$OR = \frac{\pi(1)/[1 - \pi(1)]}{\pi(0)/[1 - \pi(0)]}.$$

Vetosuhde ilmaisee, kuinka paljon todennäköisemmin tapahtuma Y toteutuu ($y=1$) ryhmässä, jossa $x=1$ verrattuna ryhmään, jossa $x=0$. Jos $OR = 1$, ei muuttujalla X ole vaikutusta vasteeseen. Jos $OR > 1$, muuttujan X olemassaolo lisää todennäköisyyttä, että $y=1$. Alle yhden olevat arvot puolestaan tarkoittavat, että muuttujan X olemassaolo vähentää y :n tapahtumisen todennäköisyyttä.

Vetosuhde saadaan laskettua kätevästi myös tarkasteltavan selittäjän kertoimen avulla, sillä voidaan osoittaa taulukon 5.1. arvojen avulla, että

$$\begin{aligned} OR &= \frac{\pi(1)/[1 - \pi(1)]}{\pi(0)/[1 - \pi(0)]} \\ &= \frac{\left(\frac{1}{1+e^{-(\beta_0+\beta_1)}} \right) \left(\frac{e^{-\beta_0}}{1+e^{-\beta_0}} \right)}{\left(\frac{1}{1+e^{-\beta_0}} \right) \left(\frac{e^{-(\beta_0+\beta_1)}}{1+e^{-(\beta_0+\beta_1)}} \right)} \\ &= \frac{e^{-\beta_0}}{e^{-(\beta_0+\beta_1)}} \\ &= e^{\beta_1}. \end{aligned}$$

Vetosuhteelle voidaan määritellä luottamusväli edellä määriteltyjen regressiokertoimien luottamusvälien avulla siten, että $100 \times (1 - \alpha)\%$ luottamusvälin alaraja on

$$e^{LV_a} = \exp[\hat{\beta}_i - z_{1-\alpha/2} * \hat{SE}(\hat{\beta}_i)]$$

ja yläraja

$$e^{LV_y} = \exp[\hat{\beta}_i + z_{1-\alpha/2} * \hat{SE}(\hat{\beta}_i)].$$

Luottamusväli on hyvä muodostaa, koska vetosuhteen jakauma on usein oikealle vino.

Vetosuhdetta käytetään muuttujien vaikutusten arvioimiseen etenkin terveystieteellisissä aineistoissa. Selitettävänä muuttujana voisi olla esimerkiksi keuhkosityövän olemassaolo, jolle voidaan etsiä riskitekijöitä selittäjien joukosta etsimällä korkeita OR -arvoja. Esimerkiksi dikotominen muuttuja, joka ilmaisee tupakoiko henkilö vai ei, voisi saada vetosuhteen arvoksi 2. Tällöin voitaisiin sanoa, että keuhkosityöpää ilmenee kaksi kertaa enemmän tupakoijien joukossa verrattuna niihin, jotka eivät polta.

Selittäjänä voi olla myös jatkuva muuttuja, kuten ikä. Jos iän vetosuhte on esimerkiksi 1,1, tarkoittaa se sitä, että vuoden lisäys iässä kasvattaa keuhkosityövän olemassaolon riskiä 1,1-kertaiseksi. Jos halutaan tietää, kuinka paljon useamman yksikön muutos kasvattaa riskiä, saadaan se kaavasta

$$OR(c) = e^{c\beta_1},$$

missä c on haluttu muutos selittäjän yksikköinä ja β_1 selittäjää vastaava regressiokerroin. Jos esimerkiksi iälle on saatu kerroinestimaatti 0,1 ja halutaan saada selville, kuinka paljon 10 vuoden lisäys iässä kasvattaa riskiä, niin edellä olevan kaavan perusteella

$$\hat{OR}(10) = e^{10 \cdot 0,1} = 2,7.$$

Henkilöllä, joka on 10 vuotta vanhempi, olisi siis 2,7-kertainen riski sairastaa keuhkosityöpää verrattuna nuorempaan. Tässä oletetaan, että iän vaikutus vasteeseen on lineaarinen, eli riski suurenee yhtä paljon riippumatta siitä, vertailaanko esimerkiksi 50- ja 60-vuotiaita vai 70- ja 80-vuotiaita henkilöitä.

5.5 Mallin valinta

Yleisesti ottaen tavoitteena on valita sellaiset muuttujat regressiomalliin, että ne selittävät mahdollisimman hyvin vastemuuttujaa. Lisäksi malliin on usein syytä valita sellaisia kovariaatteja, jotka liittyvät olennaisesti tutkittavaan ilmiöön, vaikka ne eivät olisikaan tilastollisesti merkitseviä. Selittäjien määrä tulisi kuitenkin pitää tarpeeksi pienenä, jotta malli tuottaisi numeerisesti vakaita estimaatteja. Mitä enemmän muuttujia mallissa on, sitä suurempia ovat estimoidut keskivirheet ja sitä riippuvaisemmaksi malli tulee havaitusta datasta.

Yksittäisten muuttujien merkitsevyyden arvioimisen lisäksi voidaan mallin hyvyttä arvioida esimerkiksi siten, että tutkitaan paraneeko malli oleellisesti, kun siihen lisätään selittäjiä yksi kerrallaan. Tarkastellaan seuraavaksi keinoja, joilla voidaan arvioida mallin hyvyttä ja vertailla malleja.

5.5.1 Devianssi

Mallin hyvyttä voidaan kuvata sen devianssin (deviance) avulla. Devianssi on vastaava käsite kuin jäännösneliösumma lineaarisessa regressiossa eli se kuvaa

mallin antamien ennusteiden poikkeamaa havaituista arvoista. Malli on sitä parempi, mitä pienempi devianssi on. Kun malliin lisätään muuttuja, voidaan tutkia pieneneekö devianssi selittäjän lisäämisen myötä tilastollisesti merkitsevästi. Devianssi voidaan laskea esimerkiksi siten, että

$$D = -2 \sum_{i=1}^n \left[y_i \ln \left(\frac{\hat{\pi}_i}{y_i} \right) + (1 - y_i) \ln \left(\frac{1 - \hat{\pi}_i}{1 - y_i} \right) \right],$$

missä y_i viittaa vastemuuttujan havaittuun arvoon ja $\hat{\pi}_i$ sen ennustettuun arvoon, kun estimoinnissa on käytetty suurimman uskottavuuden menetelmää.

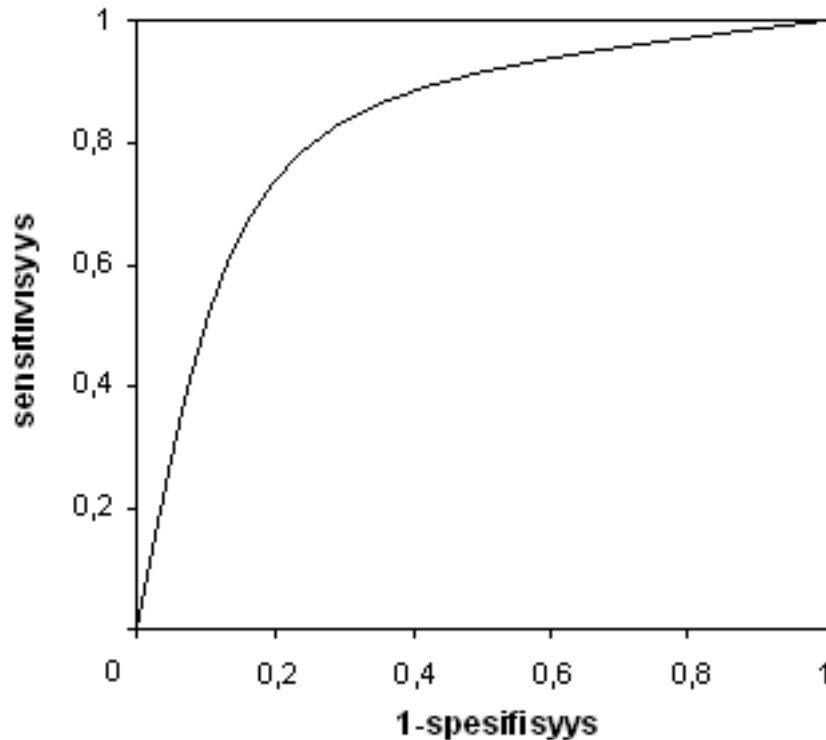
Muuttujan merkitsevyyttä mallissa voidaan testata siten, että muodostetaan kaksi mallia, suppeampi ja laajempi malli. Suppeammassa mallissa voisi olla selittäjinä vakio ja ikä ja laajemmassa mallissa lisäksi sukupuoli. Tarkoituksena on selvittää, onko sukupuoli merkitsevä selittäjä. Lasketaan kummallekin mallille devianssi. Devianssien erotuksen merkitsevyyttä voidaan testata siten, että muodostetaan testisuure $G = D(\text{suppeammalle mallille}) - D(\text{laajemmalle mallille})$. Testisuure noudattaa χ^2 -jakaumaa vapausastein p , missä p on mallien parametrien erotus (tässä tapauksessa siis 1). Tärkein testille asetettava vaatimus on riittävän suuri otoskoko. Nollahypoteesina on, että mallien devianssit ovat yhtä suuria ja vaihtoehtoisena hypoteesina, että devianssit ovat eri suuria. Jos testisuureta vastaava p -arvo on riittävän pieni, niin H_0 hylätään ja voidaan sanoa, että mallien devianssit eroavat toisistaan merkitsevästi, eli sukupuolimuuttujan mukanaolo parantaa mallia.

5.5.2 Mallin erottelukyky

Logistisessa regressiossa tarkastellaan usein mallin erottelukykä. Se tarkoittaa sitä, kuinka hyvin malli luokittelee tilastoyksiköt oikeisiin luokkiin. Erottelukykyä voidaan kuvata niin sanotun c -statistiikan (c -statistic) avulla. Tarkastellaan esimerkkinä tilannetta, jossa muodostetaan potilaista pareja, joissa toinen kuolee ja toinen jää eloon. Tunnusluku c ilmaisee niiden parien osuuden, joissa kuolleen potilaan kuoleman ennustettu todennäköisyys on suurempi kuin eloon jääneen.

C -arvo vastaa myös niin sanotun ROC -käyrän (Receiver Operating Characteristic) alapuolella olevaa aluetta (pinta-alaa). Tätä tilannetta on havainnollistettu kuviossa 5.2. Kuviossa on käytetty käsitteitä sensitiivisyys ja spesifisyys. Ennustearvojen perusteella kunkin potilaan kohdalla voidaan ennustaa, onko potilas kuollut vai elossa. Sensitiivisyys tarkoittaa oikein ennustettujen kuolleiden potilaiden osuutta kaikista kuolleista. Spesifisyys taas ilmaisee oikein ennustettujen elossa olevien potilaiden osuutta kaikista elossa olevista. Kuviossa 5.2. ROC -käyrä kuvaa käsitteiden sensitiivisyys ja (1-spesifisyys) välistä suhdetta. Käyrän alapuolelle jäävä pinta-ala siis ilmaisee c -arvon suuruuden.

C -arvot vaihtelevat lukujen 0 ja 1 välillä. Mitä lähempänä arvo on ykköstä, sitä parempi on mallin erottelukyky. Arvo 0 on sinänsä yhtä hyvä kuin arvo 1, mutta tällöin malli on väärin päin määriteltä. Arvo 0,5 vastaa ennusteky-



Kuvio 5.2. ROC -käyrä.

vyltään kolikonheittoa. (Iezzoni 2003, s. 253.) C-statistiikkaa voidaan hyödyntää esimerkiksi siten, että tutkitaan kasvattaako selittäjän lisääminen mallin c-arvoa, eli parantaako muuttujan lisääminen mallin erottelukykyä. C-arvo on selkeä mittari mallin hyvyydelle, sillä se tiivistää tiedon yhteen lukuun ja arvojen vertailu on helppoa.

C-arvo voidaan määritellä kätevästi järjestystunnuslukujen avulla. Tarkastellaan esimerkiksi potilasryhmää, jossa tietyssä ajankohdassa n_0 henkilöä elää ja n_1 henkilöä on kuollut. Kullekin potilaalle voidaan laskea muodostetun mallin perusteella ennuste kuoleman todennäköisyydelle. Näiden ennustearvojen perusteella potilaille annetaan järjestystunnusluvut siten, että pienimmän ennustearvon omaava potilas saa luvun 1, seuraava luvun 2 ja lopulta suurimman ennustearvon omaava potilas saa järjestystunnusluvun $n_0 + n_1$, joka on siis potilaiden yhteismäärä. C-arvo saadaan kaavasta

$$(5.4) \quad c = \frac{S - S_{min}}{S_{max} - S_{min}}.$$

Kaavassa (5.4) S tarkoittaa järjestystunnuslukujen summaa kuolleille potilaille (n_1 kpl). S saa pienimmän mahdollisen arvon, kun kaikki kuolleet henkilöt omaavat pienemmän ennustearvon kuin kaikki elossa olevat henkilöt. Tällöin $S_{min} = 1 + 2 + \dots + n_1$. Voidaan osoittaa, että $S_{min} = n_1(n_1 + 1)/2$. Tällöin mallilla on täydellinen erottelukyky, mutta se on väärin päin määritelty. Suurimman mahdollisen arvon S saa, kun kuolleilla henkilöillä on suurimmat

ja elossa olevilla vastaavasti pienimmät ennustetodennäköisyydet kuolemalle. Nyt $S_{max} = (n_0 + 1) + (n_0 + 2) + \dots + (n_0 + n_1)$. Tämä voidaan esittää myös muodossa $S_{max} = S_{min} + n_0 n_1$. Tällöin malli on oikein määritelty ja täydellisesti erotteleva. S voi saada arvoja näiden minimi- ja maksimiarvojen väliltä. Jos yhtälössä (5.4) $S = S_{min}$, $c = 0$. Jos taas $S = S_{max}$, $c = 1$. Jos S on yhtä lähellä minimi- ja maksimiarvoaan, c saa arvon 0,5. Tällöin järjestystunnuslukujen keskimääräinen arvo on sama sekä kuolleiden että elävien ryhmässä. Nyt mallilla ei siis ole minkäänlaista erottelukykä. (Iezzoni 2003, s. 255-256.)

C-arvolle voidaan määritellä luottamusväli esimerkiksi siten, että muodostetaan suuri määrä alkuperäisen aineiston kokoista bootstrap-otosta (otos palauttaen) ja lasketaan kustakin otoksesta c-arvo. Sen jälkeen järjestetään nämä luvut suuruusjärjestykseen. Esimerkiksi 95 prosentin luottamusväli muodostetaan siten, että poimitaan alku- ja loppupäästä sellaiset havainnot luottamusvälin ala- ja ylärajoiksi, että näiden havaintojen väliin jää 95 % havainnoista ja siis sekä alku- että loppupäähän jää 2,5 % havainnoista.

6 Tulokset

Aineiston käsittelyn ja analysoimisen olen tehnyt Survo MM -ohjelmalla. Logistisen regression tulosten ajamiseen olen käyttänyt Survon logreg -sukroa. Tutkielman teossa suuren osan ajasta veivät aineiston käsittely ja logistisen regression tulosten ajaminen ja kirjaaminen. Lähtökohtana minulla oli PERFECT -vertailuaineisto, johon tein vielä joitakin muokkauksia.

Analyyseissa käytetään vasteena 365 ja 90 päivän kuolleisuuksia. Kuten alussa mainittiin, 90 päivän kuolleisuuden voisi ajatella kuvaavan paremmin lonkkamurtumasta johtuvaa kuolleisuutta. Kuolleisuusajana voitaisiin käyttää myös esimerkiksi 30 päivää, mutta se ilmaisisi lähinnä heikkokuntoisimpien potilaiden kuolleisuutta. Nämä potilaat olisivat todennäköisesti kuolleet joka tapauksessa lähiaikoina. Lisäksi näitä tapahtumia olisi turhan vähän.

Muodostetaan aluksi molemmille kuolleisuusajoille niin sanotut perusmallit, joihin otetaan mukaan tietyt taustamuuttujat. Näihin malleihin sitten lisätään komorbiditeetteja. Tarkoituksena on tutkia, miten sairauksien lisäys vaikuttaa mallien c-arvoihin ja miten vaikutukset eroavat, kun sairauksia tarkastellaan murtumahetkestä taaksepäin eri pituisilla ajoilla. Sitten muodostetaan laajempia malleja, joihin otetaan mukaan useampia sairauksia ja mahdollisesti sairauksiin liittyviä interaktiomuuttujia. Tavoitteena on löytää oleelliset oheissairaudet ja saada selville riittävä seuranta-aika sairauksille, eli aikapiste, jota kauemmaksi menemällä ei mallin erottelukyky enää parane merkittävästi. Malleihin pyritään valitsemaan ainoastaan oleelliset sairaudet eli jätetään pois sellaisia sairauksia, joilla ei ole merkittävää vaikutusta mallin erottelukyvyn kannalta. Kuten aikaisemmin todettiin, tarkastellaan pelkästään omatoimisia potilaita.

Oleellisten liitännäissairauksien valinta ei ole yksiselitteistä ja valintaan voidaan käyttää monia kriteerejä. Tässä tutkielmassa valinta pohjautuu pääasiassa tilastollisiin tuloksiin, erityisesti c-arvojen tarkasteluun. Sairauksien valinnassa on kuitenkin syytä kiinnittää huomiota myös sairauden yleisyyteen ja diagnoosien luotettavuuteen. Lisäksi voi olla tarpeellista valita sellaisia sairauksia, joilla tiedetään olevan vaikutusta hoidon kannalta, vaikka ne eivät olisikaan tilastollisesti merkitseviä.

Seuraavissa alaluvuissa käytetään vasteena ensin vuoden kuolleisuutta ja sen jälkeen tarkastellaan, millä tavalla tulokset muuttuvat, kun vasteena on kolmen kuukauden kuolleisuus. Tutkielman liitteenä on perusmallien logistisen regression tulokset sekä tärkeimpiin oheissairauksiin liittyviä tilastoja kummallekin kuolleisuusajalle.

6.1 365 päivän kuolleisuus

Muodostetaan aluksi malli, jossa selitettävänä muuttujana on vuoden kuolleisuus (onko potilas kuollut vai elossa vuoden kuluttua indeksihoitojakson alusta). Selittäjinä on vakio, sukupuoli, ikä, murtumatyyppi sekä muuttuja, joka kertoo onko potilas ollut lyhytaikaisessa laitoshoidossa murtumaa edeltäneiden kolmen kuukauden aikana.

Mallin tulosteista nähdään (ks. liite A), että sukupuoli vaikuttaa kuolleisuuteen siten, että miehillä on suurempi todennäköisyys kuolla kuin naisilla. Iän vaikutus on puolestaan sellainen, että vanhemmilla potilailla on suurempi kuolleisuusriski kuin nuoremmilla. Tässä iän vaikutus oletetaan lineaariseksi. Kun iän vaikutusta tarkasteltiin käyttämällä eri potensseja, havaittiin että iän kolmas potenssi selittää kuolleisuutta hieman paremmin. Käytetään kuitenkin jatkossa lineaarista ikämuuttujaa, sillä sen tulkinta on selkeämpää. Potilailla, jotka olivat olleet lyhytaikaisessa laitoshoidossa murtumaa edeltäneiden kolmen kuukauden aikana, on tulosten perusteella yli kaksinkertainen riski kuolla ($OR=2,34$). Lyhytaikaista laitoshoidoa oli saanut 34 % aineiston omatoimisista potilaista. Edellä mainitut muuttujat olivat selvästi tilastollisesti merkitseviä (p-arvot likimain 0). Myös murtumatyyppi oli merkitsevä siten, että potilailla, joilla oli todettu subtrokanteerinen murtuma, on suurempi kuolleisuusriski verrattuna muihin murtumatyyppeihin. Pidetään siis kaikki edellä mainitut muuttujat mallissa.

Tähän perusmalliin lisättiin jokainen liitännäissairaus erikseen yksi kerrallaan. Yleisimmät sairausryhmät (10 vuoden ajalla) olivat verenpainetauti, masennus ja muut vammat. Komorbiditeettia AIDS/HIV sairastaneita oli vain muutama, joten tämä sairausryhmä jätetään saman tien pois jatkotarkasteluista. Joidenkin sairauksien kohdalla (kuten ylipaino ja laihtuminen) kaikkia tapauksia ei varmastikaan ole diagnosoitu, joten tällaisten sairauksien tapauksessa tulokset eivät ole kovin luotettavia.

Vähintään yhdessä aikapisteessä tilastollisesti merkitseviä (p-arvo $< 0,05$) liitännäissairauksia löytyi 24 kappaletta. Lukuun ottamatta komorbiditeettia muut vammat, sairauksien esiintyminen lisäsi kuoleman todennäköisyyttä. Perusmallin c-arvo (ilman sairauksia) oli 0,706. C-arvo nousi tästä käytännössä ainoastaan silloin, kun sairaus oli selvästi tilastollisesti merkitsevä (p-arvo likimain 0). Läheskään kaikki komorbiditeetit eivät kuitenkaan nostaneet c-arvoa.

Oheissairauksia tarkasteltiin eri pituisilla ajoilla taaksepäin murtumasta, mutta suurin mielenkiinto kohdistuu tässä vaiheessa kuitenkin olennaisten sairauksien valitsemiseen. Useimpien sairauksien kohdalla c-arvot kasvoivat hieman tai pysyivät samoina, kun tarkasteluaikaa pidennettiin, mutta esimerkiksi syövän tapauksessa arvot pienenevät hieman. C-arvojen vaihtelu oli kuitenkin varsin vähäistä aikapisteiden välillä. Keskimäärin korkeimmat c-arvot saatiin, kun sairauksia tarkasteltiin kolme vuotta taaksepäin. Käytetään jatkossa tarkasteluaikana kolmen vuoden kohtaa ja tutkitaan myöhemmin, onko tarkasteluajan pituudella merkitystä vai riittäisikö lyhyempikin tarkasteluaika.

Voimakkaimmin kuolleisuusriskiä näyttivät lisäävän syöpä ja sydämen va-

jaatoiminta. Nämä sairausryhmät erottuivat selvästi muista sairauksista, kun tarkasteltiin c-arvoja, joten ne on ainakin syytä ottaa mukaan selittäjiksi. Lisäksi malliin valitaan neljä sellaista komorbiditeettia (psykoosit, keuhkokuume, diabetes ja perifeerinen valtimotauti), joiden lisääminen kasvattaa mallin c-arvoa vielä jonkin verran.

Taulukossa 6.1. on esitetty kussakin vaiheessa malliin lisättävät muuttujat sekä kyseisen mallin c-arvo. Lisäksi taulukossa on tarkasteltu mallien devianssien erotuksia, jotta voidaan arvioida, parantaako kunkin muuttujan lisääminen mallia merkitsevästi. Erotukset noudattavat χ^2 -jakaumaa vapausastein 1. Mitä enemmän devianssi pienenee, sitä paremmaksi malli tulee. Muuttujat on lisätty siinä järjestyksessä, että lisätään aina eniten devianssia pienentävä muuttuja tarkasteltavien muuttujien joukosta. Havaittiin, että kaikkien lisättyjen muuttujien kohdalla devianssien erotus oli selvästi tilastollisesti merkitsevää eli muuttujien lisääminen paransi mallia. Devianssi pieneni merkitsevästi muutamien muunkin oheissairausten kohdalla, mutta muutos oli jo pienempi eikä c-arvo enää kasvanut juurikaan. C-arvo toki kasvaisi vielä, jos malliin lisättäisiin useampi sairaus kerralla, mutta nyt halutaan valita ainoastaan tärkeimmät oheissairaudet.

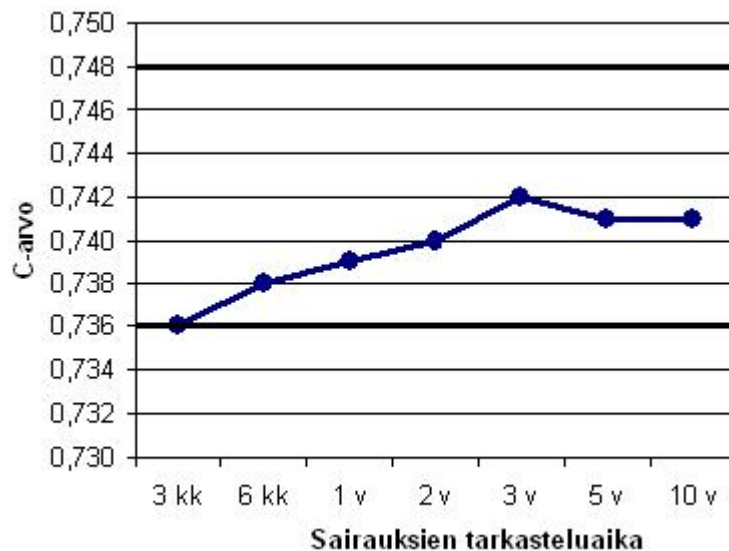
Taulukko 6.1. C-arvojen ja devianssien muutos lisättäessä selittäjiä kumulatiivisesti perusmalliin, kun vasteena on vuoden kuolleisuus. P-arvo on kaikkien lisäyksien kohdalla likimain 0.

Lisätyt muuttujat	c-arvo	devianssin muutos
perusmalli	0,706	28 203
+syöpä	0,719	-358
+sydämen vajaatoiminta	0,730	-332
+psykoosit	0,733	-104
+keuhkokuume	0,735	-57
+diabetes	0,736	-30
+perifeerinen valtimotauti	0,737	-27
+syövän ja iän interaktio	0,741	-167
+syövän ja sukupuolen interaktio	0,742	-33

Seuraavaksi tarkasteltiin komorbiditeettien ja iän sekä sukupuolen välisiä yhdysvaikutuksia eli interaktioita. Monella sairaudella (mm. keuhkojen veren-

kierron häiriöt, krooniset keuhkosairaudet ja syöpä) oli pienempi vaikutus kuolleisuuden vanhemmilla potilailla. Sukupuolen vaikutus oli tilastollisesti merkitsevää muutaman oheissairauden (erityisesti rytmihäiriöt ja syöpä) tapauksessa siten, että sairauden vaikutus kuolleisuuteen oli suurempi naisilla. Osoittautui kuitenkin, että ainoastaan syövän interaktiot iän ja sukupuolen kanssa nostivat c-arvoja selvästi. Tämän vuoksi muita interaktiotermejä ei oteta mukaan, vaikka osa niistä olikin merkitseviä.

Päädytään lopulliseen malliin taulukon 6.1. askelten mukaisesti. Tässä mallissa kuolleisuutta selittää yhteensä 13 muuttujaa (vakio, 4 taustamuuttujaa, 6 liitännäissairautta ja 2 interaktiomuuttujaa). Mallin c-arvo on 0,742, kun sairauksia on tarkasteltu kolme vuotta taaksepäin murtumasta. Havaitaan, että tämä on korkein arvo, kun sairauksia tarkastellaan taaksepäin murtumasta eri pituisilla ajoilla. C-arvot eivät kuitenkaan ole juurikaan pienempiä muissakaan aikapisteissä (kuvio 6.1.). Tutkitaan seuraavaksi, ovatko nämä erot merkitseviä.



Kuvio 6.1. C-arvojen muutos eri tarkasteluajoilla, kun vasteena on vuoden kuolleisuus. Kuvioon on lisäksi merkitty paksunnetuilla viivoilla 95 % luottamusvälin rajat arvoille, jotka on laskettu mallista, jossa sairauksia on tarkasteltu kolme vuotta taaksepäin.

Muodostetaan 500 bootstrap-otosta datasta, jossa sairauksia on tarkasteltu kolme vuotta murtumasta taaksepäin. Kustakin otoksesta lasketaan c-arvo ja näin saadaan arvoille 95 % luottamusväli, jonka rajat on merkitty kuvioon 6.1. Luottamusvälin avulla voidaan nyt arvioida sitä, eroaako kolmen vuoden huippukohta merkitsevästi muista arvoista. Nähdään, että kolmen kuukauden tarkasteluajalla havaittu arvo on likimain sama kuin luottamusvälin alaraja, muut arvot mahtuvat selvemmin luottamusvälille. Vaikka arvot siis kasvavat jonkin verran tarkastelu-aikaa pidennettäessä, ero pienimmän ja suurimman

arvonkaan välillä ei ole merkitsevää. Toisaalta luottamusvälin leveydestä voidaan päätellä, että pienet erot c -arvoissa eivät ole merkitseviä.

Edellä käytettiin tarkasteluaikana kolmea vuotta. Lyhyempikin aika kuitenkin näyttää riittävän, sillä on turhaa kerätä hoitotietoja niin pitkältä ajalta, jos malli ei enää parane. Kolme ja kuusi kuukautta tuntuvat kuitenkin liian lyhyiltä ajoilta, koska sairaushavainnot ei näillä ajoilla kerry niin paljon. Toisaalta uusia havainnot tulee jo vähemmän suhteessa aikaan, kun tarkasteluaikaa pidennetään vuodesta eteenpäin. Mitään tiettyä aikaa ei kuitenkaan pystytä valitsemaan tilastollisin perustein, vaan käytettävä seuranta-aika voidaan valita esimerkiksi sen mukaan, kuinka aikaavievää tai hankalaa sairaustietojen kerääminen on.

6.2 90 päivän kuolleisuus

Kolmen kuukauden mallin lähtökohtana ovat samat selittäjät kuin edellä, ainoastaan vastemuuttujia on siis eri (onko potilas kuollut vai elossa kolmen kuukauden kuluttua indeksijakson alusta). Liitteessä A on myös tämän perusmallin tärkeimmät tulosteet. Kertoimien estimaateissa ja merkitsevyyksissä ei ole suuria eroja, kaikki taustamuuttujat ovat nytkin merkitseviä. Enemmänkin olemme kiinnostuneita liitännäissairauksien vaikutusten eroista, kun kuolleisuus aika on lyhyempi. Tähän malliin siis lisättiin vastaavasti kukin sairaus vuorotellen. Mallin c -arvo ilman sairauksia oli nyt 0,721 eli jonkin verran suurempi kuin edellä vuoden kuolleisuudella (0,706). Vasteena kokeiltiin käyttäen myös 30 ja 120 päivän kuolleisuuksia, mutta näiden mallien c -arvot olivat hieman pienempiä. Vaikuttaa siis, että nimenomaan 90 päivän kuolleisuus olisi hyvä vaste tämän aineiston tapauksessa.

Eniten c -arvoa nostivat nytkin sydämen vajaatoiminta ja syöpä. Muutenkin sairauksien vaikutus kuolleisuuteen oli samansuuntaista kuin edellä. Suurin ero sairauksien merkitsevyyksissä oli alkoholin väärinkäytön kohdalla siten, että 365 päivän kuolleisuutta käytettäessä se oli jonkin verran merkitsevä, mutta ei lainkaan merkitsevä 90 päivän kuolleisuudella. Tilastollisesti merkitseviä oheissairauksia löytyi 22 kappaletta. Merkitsevät sairaudet olivat joiltakin osin erilaisia verrattuna vuoden kuolleisuuden tilanteeseen, jossa merkitseviä sairauksia oli siis 24 kappaletta. C -arvojen vaihtelu aikapisteiden välillä oli nytkin varsin pientä. Käytetään tarkasteluaikana tässäkin kolmea vuotta, kun valitaan oleellisia tekijöitä ja tutkitaan lopuksi riittäkö lyhyempikin aika.

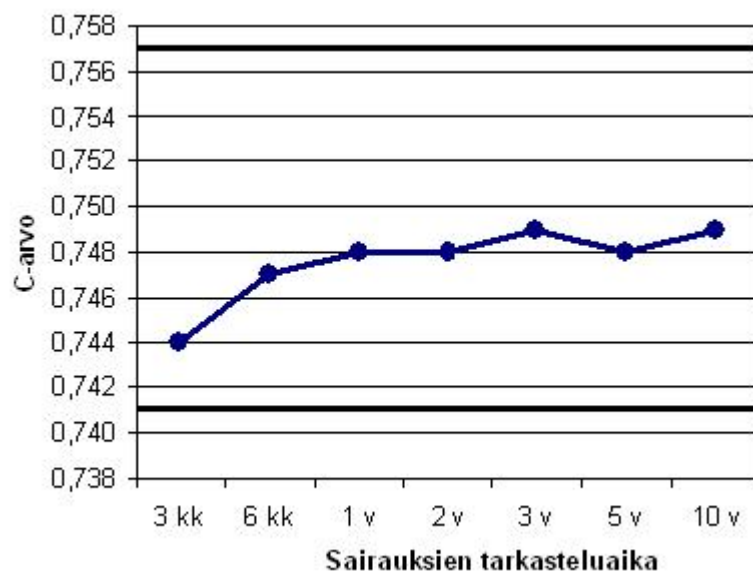
Nytkin itsestään selvät valinnat malliin ovat syöpä ja sydämen vajaatoiminta. Lisäksi liitännäissairauksista otetaan mukaan psykoosit ja keuhkokuume. Interaktiomuuttujien vaikutukset olivat melko lähellä vuoden kuolleisuudella saatuja tuloksia. Nytkin syövän interaktiot olivat merkitsevimmät, mutta ainoastaan syövän ja iän välinen interaktiotermi on nyt tarpeellista ottaa mukaan malliin. Taulukossa 6.2. on esitetty muuttujien lisäämisprosessi ja c -arvojen sekä devianssien muutokset. Devianssien erotukset olivat myös nyt selvästi merkitseviä jokaisen lisätyn selittäjän kohdalla. Nytkin malliin on valittu

ainoastaan merkitsevimmät tekijät.

Taulukko 6.2. C-arvojen ja devianssien muutos lisättäessä selittäjiä kumulatiivisesti perusmalliin, kun vasteena on 90 päivän kuolleisuus. P-arvo on kaikkien lisäyksien kohdalla likimain 0.

Lisätyt muuttujat	c-arvo	devianssin muutos
perusmalli	0,721	19 968
+sydämen vajaatoiminta	0,732	-216
+syöpä	0,740	-140
+psykoosit	0,743	-62
+keuhkokuume	0,745	-30
+syövän ja iän välinen interaktio	0,749	-131

Lopulliseen malliin valitaan nyt yhteensä 10 selittäjää (vakio, 4 taustamuuttujaa, 4 liitännäissairautta ja 1 interaktiomuuttuja). Tämä malli saa korkeimmat c-arvot, kun sairauksia tarkastellaan 3 tai 10 vuotta taaksepäin murtumasta. Kuten kuviosta 6.2. nähdään, erot ovat nytkin pieniä eri aikapisteiden välillä. Muodostetaan 95 % luottamusväli samoin kuin edellä mallille, jossa sairauksien tarkasteluajana on kolme vuotta. Kuviosta 6.2. nähdään, että kaikki havaitut arvot mahtuvat selvästi tälle välille eli suurimman ja pienimmän arvonkaan ero ei ole merkitsevä. Ei siis välttämättä tarvitse tarkastella sairauksia kolmea vuotta taaksepäin, sillä erot c-arvoissa aikapisteiden välillä eivät ole merkitseviä. Myöskään 90 päivän kuolleisuusaikaa käytettäessä ei siis pystytä valitsemaan yhtä tiettyä tarkasteluajaa.



Kuvio 6.2. C-arvojen muutos eri tarkasteluajoilla, kun vasteena on 90 päivän kuolleisuus. Kuvioon on lisäksi merkitty paksunnetuilla viivoilla 95 % luottamusvälin rajat arvoille, jotka on laskettu mallista, jossa sairauksia on tarkasteltu kolme vuotta taaksepäin.

7 Yhteenveto

Tutkielmassa oli tavoitteena etsiä sellaisia tekijöitä, jotka tulisi vakioida, kun halutaan vertailla luotettavasti lonkkamurtumapotilaiden hoitokäytäntöjä. Tilastollisesti merkitseviksi taustamuuttujiksi osoittautuivat ikä, sukupuoli, murtumatyyppi sekä muuttuja, joka ilmaisee onko potilas saanut lyhytaikaista laitoshoidoa murtumaa edeltäneiden kolmen kuukauden aikana. Kuolleisuusriski näyttivät kasvattavan korkea ikä, miessukupuoli, subtrokanteerinen murtumatyyppi sekä lyhytaikaisessa laitoshoidossa olo. On kuitenkin syytä huomioida, että vakioitaviksi tekijöiksi voitiin valita ainoastaan sellaisia muuttujia, jotka löytyvät rekistereistä. Vakiointia ei siis välttämättä pystytä tekemään kaikkien haluttujen tekijöiden suhteen.

Erityisen mielenkiinnon kohteena olivat potilailla havaitut muut sairaudet. Sairauksien vaikutusta kuolleisuuteen tarkasteltiin käyttämällä vasteena kah- ta eri pituista kuolleisuusmuuttujaa. Havaittiin, että 90 päivän kuolleisuutta käytettäessä mallin erottelukyky oli parempi kuin 365 päivän kuolleisuuden tapauksessa. Tuloksissa ilmeni jonkin verran eroja sairauksien merkitsevyyk- sissä eri kuolleisuusajoilla, mutta merkitseviä sairausryhmiä löytyi kummalla- kin kuolleisuusajalla yli 20. Lopullisiin malleihin valittiin kuitenkin ainoastaan tärkeimmät oheissairaudet. Havaittiin, että sairauksista ainakin syöpä ja sydä- men vajaatoiminta tulisi huomioida potilasaineksen vakioinnissa. Muiden sai- rauksien valinnan osalta voidaan päätyä erilaisiin lopputuloksiin riippuen siitä, kuinka monta komorbiditettia halutaan valita ja minkälaisia valintakriteerejä käytetään.

Oheissairauksien vaikutusta kuolleisuuteen tarkasteltiin eri pituisilla ajoilla murtumaa edeltäneeltä ajalta. Tavoitteena oli löytää lyhin riittävä seuranta- aika sairauksille. Havaittiin, että sairauksien vaikutusten erot olivat kuitenkin pieniä aikapisteiden välillä ja mitään yhtä tiettyä aikaa ei pystytty tilastollisin perustein valitsemaan. Todettiin, että kolme ja kuusi kuukautta tuntuvat liian lyhyiltä seuranta-ajoilta, koska sairaushavainnot ei kerry niin paljon näillä ajoilla. Toisaalta on turhaa kerätä aineistoa kovin pitkältä ajalta, koska mallin erottelukyky ei parantunut, vaikka tarkasteluaikaa pidennettiin. Suurin c-arvo saatiin pääsääntöisesti kolmen vuoden kohdalla. Koska erot eivät kuitenkaan olleet merkitseviä, voidaan jo esimerkiksi yhden vuoden tarkasteluaikaa pitää riittävänä.

Tässä tutkielmassa esitettyjä tuloksia ja menetelmiä voitaisiin käyttää myös silloin, kun etsitään muihin sairausryhmiin liittyviä oleellisia liitännäissairauk- sia. Voitaisiin kokeilla käyttää myös muita tilastollisia menetelmiä ja vaste-

muuttujia vakioitavien tekijöiden valinnassa.

Haluaisin lopuksi kiittää FM Antti Liskiä ja VTT Reijo Sundia aiheeseen perehdyttämisestä ja työn ohjauksesta. Lisäksi haluan kiittää Stakesia, joka luovutti aineiston käyttöni tutkimuksen ajaksi.

Lähdeluettelo

- Duodecim (2006), ”Käypä hoito -suositus. Lonkkamurtumapotilaiden hoito.”, *Duodecim* 122(3):358-79
- Holman, C.D., Preen, D.B., Baynham, N.J., Finn, J.C. & Semmens, J.B. (2005), ”A multipurpose comorbidity scoring system performed better than the Charlson index”, *Journal of Clinical Epidemiology*, 58, 1006–1014.
- Hosmer, D. W. & Lemeshow S. (1989), ”Applied Logistic Regression”, *Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, New York*
- Iezzoni, L.I. (2003), ”Risk Adjustment for Measuring Health Care Outcomes”, *Chicago: Health Administration Press*
- Jiang, H.X., Majumdar, S.R., Dick, D.A., Moreau, M., Raso, J., Otto, D.D. & Johnston D.W. (2005), ”Development and Initial Validation of a Risk Score for Predicting In-Hospital and 1-Year Mortality in Patients With Hip Fractures”, *Journal of Bone Miner and Mineral Research*, 20(3), 494–500.
- Juntunen, M., Sund, R., Peltola M., & Häkkinen U. (2008), ”Potilasrakenteen erojen huomioon ottaminen erikoissairaanhoidon vaikuttavuuden rekisteritutkimuksessa”, *Sosiaalilääketieteellinen aikakauslehti 2008:45:258–272*
- Sund, R. (2000), ”Tilastollisia menetelmiä dynaamisten potilaspopulaatioiden mallintamiseen. Tapahtumahistoria-analyysia hoitoilmoitusrekisterin skitsofreeni-koille”, *Aiheita 26/2000, Stakes, Helsinki*
- Sund, R. & Liski A. (2005), ”Quality Effects of Operative Delay on Mortality in the Case of Hip Fracture Treatment”, *Quality and safety in health care*, 14, 371-377.
- Sund, R., Juntunen M., Lüthje P., Huusko T., Mäkelä M., Linna, M., Liski A. & Häkkinen U. (2008), ”Hoitoketjujen toimivuus, vaikuttavuus ja kustannukset lonkkamurtumapotilailla”, *Stakes, Työpapereita 18/2008, Helsinki 2008*
- Quan, H., Sundararajan, V., Halfon, P., Fong, A., Burnand, B., Luthi, J.C., Saunders, L.D., Beck, C.A., Feasby, T.E. & Ghali, W.A. (2005), ”Coding Algorithms for Defining Comorbidities in ICD-9-CM and ICD-10 Administrative Data”, *Medical Care*, 43(11), 1130–1139.

Liite A

Logistisen regression tulosteet

Taulukoissa A.1. ja A.2. on logistisen regression tulokset ns. perusmalleille kummallakin kuolleisuusajalla. Kunkin selittäjän kohdalla on estimaattiarvot sen kertoimelle ja keskivirheelle, Waldin testisuureen arvo, p-arvo sekä vetosuhteen arvo (OR) ja sen 95 % luottamusvälin rajat. Sukupuolimuuttuja on koodattu siten, että mies=0 ja nainen=1. Murtumatyyppi on otettu mukaan kahdena dummy-muttujana siten, että jos muuttuja Tyyppi2 saa arvon 1, kyseessä on trokanteerinen murtuma ja jos Tyyppi3=1, kyseessä on subtrokanteerinen murtuma. Jos molemmat muuttujat saavat arvon nolla, potilaalla on havaittu reisiluun kaulan murtuma (perustaso).

Taulukko A.1. Logistisen regression tulokset, kun vasteena on vuoden kuolleisuus.

Parametri	estimaatti	keskivirhe	Wald X^2	Wald p
Vakio	-6.232	0.144	1870	0.000000
Sukupuoli	-0.779	0.033	559	0.000000
Ikä	0.065	0.002	1283	0.000000
Tyyppi2	-0.001	0.033	0.001	0.974429
Tyyppi3	0.151	0.057	6.95	0.008378
Laitos	0.849	0.030	817	0.000000

	OR	alaraja	yläraja
Vakio	0.002	0.001	0.003
Sukupuoli	0.46	0.43	0.49
Ikä	1.07	1.06	1.07
Tyyppi2	1.00	0.94	1.06
Tyyppi3	1.16	1.04	1.30
Laitos	2.34	2.21	2.48

Taulukko A.2. Logistisen regression tulokset, kun vasteena on 90 päivän kuolleisuus.

Parametri	estimaatti	keskivirhe	Wald X^2	Wald p
Vakio	-7.676	0.191	1613	0.000000
Sukupuoli	-0.895	0.040	507	0.000000
Ikä	0.075	0.002	995	0.000000
Tyyppi2	-0.019	0.041	0.22	0.638574
Tyyppi3	0.148	0.072	4.29	0.038260
Laitos	0.816	0.037	493	0.000000

	OR	alaraja	yläraja
Vakio	0.0005	0.0003	0.0007
Sukupuoli	0.41	0.38	0.44
Ikä	1.08	1.07	1.08
Tyyppi2	0.98	0.91	1.06
Tyyppi3	1.16	1.01	1.33
Laitos	2.26	2.10	2.43

Liite B

Liitännäissairaudet ja tärkeimpiin sairauksiin liittyvät tilastot

Seuraavalla sivulla (taulukko B.1.) on esitetty kaikki tutkielmassa käytetyt liitännäissairaudet. Suluissa on englanninkielinen termi. Lisäksi sairauden perään on merkitty, jos siihen liittyy erityiskorvausoikeus (E) tai lääkekorvaus (L).

Edempänä on taulukkoja, joissa on tärkeimmät tilastot malleihin valituille oheissairauksille kummallakin kuolleisuusajalla. Taulukoissa on kunkin aikapisteen kohdalla regressiokertoimen estimaatti, Waldin testisuuretta vastaava p-arvo, vetosuhde (OR), mallin c-arvo ja kyseistä sairautta sairastaneiden lukumäärä (aineistossa potilaita yhteensä 28 797). Malleissa on käytetty selittäjinä kyseisen sairauden lisäksi valittuja taustamuuttujia (ks. liite A). Taulukot B.2. ja B.3. liittyvät 365 päivän kuolleisuuteen ja taulukko B.4. liittyy 90 päivän kuolleisuuden malliin.

Taulukko B.1. Tutkitut liitännäissairaudet.

Sydämen vajaatoiminta (Congestive heart failure) E
Rytmihäiriöt (Cardiac arrhythmias) E + L
Sydämen läppävika (Valvular disease)
Keuhkojen verenkierron häiriöt (Pulmonary circulation disorders)
Perifeerinen valtimotauti (Peripheral vascular disorders)
Verenpainetauti (Hypertension) E + L
Halvaus (Paralysis)
Parkinsonin tauti (Parkinson disease) E + L
Dementia E + L
Muut neurologiset sairaudet (Other neurological disorders)
Krooniset keuhkosairaudet (Chronic pulmonary disorders) E + L
Diabetes E + L
Kilpirauhasen vajaatoiminta (Hypothyroidism) E + L
Munuaisten vajaatoiminta (Renal failure) E
Maksasairaus (Liver disease)
Mahahaava (Peptic ulcer disease)
AIDS / HIV
Syöpä (Cancer) E + L
Nivelreuma / Sidekudostaudit E (Rheumatoid arthritis / collagen vascular diseases)
Hyytymishäiriöt (Coagulopathy)
Ylipaino (Obesity)
Laihtuminen (Weight loss)
Neste- ja elektrolyyttitasapainon häiriöt (Fluid and electrolyte disorders)
Raudanpuuteanemia (Blood loss anemia)
Puuteanemia (Deficiency anemia)
Alkoholin väärinkäyttö (Alcohol abuse)
Lääkkeiden väärinkäyttö (Drug abuse)
Psykoosit (Psychoses) E + L
Masennus (Depression) L
Keuhkokuume (Pneumonia)
Virtsatieinfektio (Urinary tract infection)
Muut vammat (Injuries)

Taulukko B.2. Tilastoja sairauksille, kun vasteena on 365 päivän kuolleisuus, osa 1/2.

Liitännäissairaus	aika	kerroin	p-arvo	OR	C-arvo	lkm
Diabetes	3 kk	0,319	0,000	1,38	0,708	3951
	6 kk	0,328	0,000	1,39	0,709	4064
	1 v	0,338	0,000	1,40	0,709	4152
	2 v	0,332	0,000	1,39	0,709	4253
	3 v	0,337	0,000	1,40	0,709	4305
	5 v	0,329	0,000	1,39	0,709	4374
	10 v	0,326	0,000	1,39	0,709	4420
Keuhkokuume	3 kk	0,397	0,000	1,49	0,707	575
	6 kk	0,403	0,000	1,50	0,707	908
	1 v	0,438	0,000	1,55	0,708	1467
	2 v	0,440	0,000	1,55	0,709	2220
	3 v	0,480	0,000	1,62	0,710	2771
	5 v	0,460	0,000	1,58	0,710	3439
	10 v	0,455	0,000	1,58	0,711	3821
Perifeerinen valtimotauti	3 kk	0,560	0,000	1,75	0,707	271
	6 kk	0,488	0,000	1,63	0,707	404
	1 v	0,426	0,000	1,53	0,707	603
	2 v	0,442	0,000	1,56	0,707	906
	3 v	0,454	0,000	1,57	0,707	1115
	5 v	0,439	0,000	1,55	0,708	1383
	10 v	0,461	0,000	1,59	0,708	1633
Psykoosit	3 kk	0,346	0,000	1,41	0,708	2917
	6 kk	0,379	0,000	1,46	0,708	3460
	1 v	0,379	0,000	1,46	0,708	3846
	2 v	0,380	0,000	1,46	0,709	4230
	3 v	0,379	0,000	1,46	0,709	4425
	5 v	0,373	0,000	1,45	0,709	4660
	10 v	0,362	0,000	1,44	0,709	4793

Taulukko B.3. Tilastoja sairauksille, kun vasteena on 365 päivän kuolleisuus, osa 2/2.

Liitännäissairaus	aika	kerroin	p-arvo	OR	C-arvo	lkm
Sydämen vajaatoiminta	3 kk	0,565	0,000	1,76	0,713	4453
	6 kk	0,578	0,000	1,78	0,714	4656
	1 v	0,603	0,000	1,83	0,715	4948
	2 v	0,620	0,000	1,86	0,716	5336
	3 v	0,624	0,000	1,87	0,717	5573
	5 v	0,638	0,000	1,89	0,718	5823
	10 v	0,652	0,000	1,92	0,719	6034
Syöpä	3 kk	1,021	0,000	2,78	0,719	1955
	6 kk	0,975	0,000	2,65	0,720	2205
	1 v	0,922	0,000	2,51	0,719	2470
	2 v	0,863	0,000	2,37	0,718	2747
	3 v	0,847	0,000	2,33	0,719	2926
	5 v	0,768	0,000	2,15	0,717	3237
	10 v	0,724	0,000	2,06	0,717	3548

Taulukko B.4. Tilastoja sairauksille, kun vasteena on 90 päivän kuolleisuus.

Liitännäissairaus	aika	kerroin	p-arvo	OR	C-arvo	lkm
Keuhkokuume	3 kk	0,344	0,001	1,41	0,721	575
	6 kk	0,356	0,000	1,43	0,722	908
	1 v	0,398	0,000	1,49	0,723	1467
	2 v	0,416	0,000	1,52	0,723	2220
	3 v	0,424	0,000	1,53	0,724	2771
	5 v	0,379	0,000	1,46	0,724	3439
	10 v	0,368	0,000	1,44	0,724	3821
Psykoosit	3 kk	0,374	0,000	1,45	0,723	2917
	6 kk	0,400	0,000	1,49	0,724	3460
	1 v	0,383	0,000	1,47	0,724	3846
	2 v	0,370	0,000	1,45	0,724	4230
	3 v	0,365	0,000	1,44	0,724	4425
	5 v	0,376	0,000	1,46	0,724	4660
	10 v	0,366	0,000	1,44	0,724	4793
Sydämen vajaatoiminta	3 kk	0,554	0,000	1,74	0,729	4453
	6 kk	0,560	0,000	1,75	0,729	4656
	1 v	0,587	0,000	1,80	0,730	4948
	2 v	0,601	0,000	1,82	0,731	5336
	3 v	0,614	0,000	1,85	0,732	5573
	5 v	0,613	0,000	1,85	0,733	5823
	10 v	0,619	0,000	1,86	0,733	6034
Syöpä	3 kk	0,749	0,000	2,11	0,729	1955
	6 kk	0,726	0,000	2,07	0,729	2205
	1 v	0,693	0,000	2,00	0,729	2470
	2 v	0,636	0,000	1,89	0,728	2747
	3 v	0,625	0,000	1,87	0,728	2926
	5 v	0,554	0,000	1,74	0,727	3237
	10 v	0,543	0,000	1,72	0,727	3548