

Aalto-yliopisto

Perustieteiden korkeakoulu

Teknillisen fysiikan ja matematiikan tutkinto-ohjelma

Epävarmuuden huomioiminen diagnostisen testausstrategian valinnassa

Kandidaatintyö

20.5.2015

Jesse Myrberg

Työn saa tallentaa ja julkistaa Aalto-yliopiston avoimilla verkkosivuilla. Muilta osin kaikki oikeudet pidätetään.



Tekijä Jesse Myrberg

Työn nimi Epävarmuuden huomioiminen diagnostisen testausstrategian valinnassa

Koulutusohjelma Teknillisen fysiikan ja matematiikan tutkinto-ohjelma

Pääaine Systeemitieteet

Pääaineen koodi F3010

Vastuopettaja Ahti Salo

Työn ohjaaja(t) Yrjänä Hynninen

Päivämäärä 20.5.2015

Sivumäärä 23 + 2

Kieli Suomi

Tiivistelmä:

Kliinikoiden vastaanotolle tulee päivittäin potilaita, joiden sairauden tilaa pyritään selvittämään potilasdiagnoosin avulla. Potilasdiagnoosissa klinikko estimoii potilaan sairauden todennäköisyyttä eli niin sanottua prioritodennäköisyyttä esimerkiksi potilaan oireiden ja sairaushistorian perusteella. Diagnoosin jälkeen klinikon on päätettävä potilaalle suoritettavista jatkotoimenpiteistä, joita tässä työssä kutsutaan testausstrategioiksi. Näitä strategioita ovat:

1. Potilaalle ei tehdä toimenpiteitä.
2. Potilaan sairauden tilasta hankitaan lisäinformaatiota diagnostisen testin avulla.
3. Potilaan hoito aloitetaan ilman lisäinformaation hankkimista.

Kaikkiin näihin testausstrategioihin liittyy epävarmuutta eri parametrien suhteen. Esimerkiksi diagnostiset testit eivät useinkaan ole täydellisiä, sillä testin antama tulos voi olla virheellinen. Myöskään testauskustannusten määrittäminen ei ole yksikäsitteistä, vaan sitä on perusteltua mallintaa intervallina eli välinä, jonka määrittävät kustannusten ala- ja ylärajat. Lisäksi hoitamiseen tai hoitamatta jättämiseen liittyviä hyötyjä terveen tai sairaan potilaan tapauksissa ei voida estimoida yksikäsitteisesti, vaan niihin liittyy epävarmuutta. Tässä työssä kullekin testausstrategialle lasketaan odotusarvoinen hyöty lääkärin estimoiman prioritodennäköisyyden funktiona, kun edellä esitettyjä parametreja mallinnetaan intervalleina tietyn perusarvon ympärillä.

Työn tavoitteena on tarkastella, kuinka parametreihin liittyvät epävarmuudet vaikuttavat odotusarvoisen hyödyn maksimoivan testausstrategian valintaan. Työssä käytetyn esimerkin perusteella on mahdollista löytää muita vaihtoehtoja parempia testausstrategioita tietyille prioritodennäköisyyden väleille, kun parametrien annetaan liikkua intervalleilla. Kaikille prioritodennäköisyyden väleille strategian valinta ei kuitenkaan ole yksikäsitteistä, sillä strategiaa voidaan tarkastella eri näkökulmista. Jos tarkastelun kohteeksi rajataan vain odotusarvoisten hyötyjen maksimit, minimi tai perusarvot, tarkasteltavat testausstrategiat voivat poiketa toisistaan merkittävästi.

Avainsanat diagnostinen, testausstrategia, epävarmuus, hyötyfunktio

Sisältö

1. Johdanto	1
2. Teoreettinen tausta	2
2.1. Prioritodennäköisyys ja sen estimointi.....	2
2.2. Posterioritodennäköisyys	3
2.3. Testin tarkkuus ja erottelukyky	4
2.4. Odotusarvoinen hyöty ja päätöspuut.....	7
3. Epävarmuuden huomioiminen	12
3.1. Perustelut epävarmuuden mallintamiselle	12
3.2. Parametrien mallinnus	12
3.3. Dominanssi	15
4. Esimerkki	15
4.1. Taustaa	15
4.2. Rajaukset	15
4.3. Testiasetelma ja tulokset.....	16
4.4. Tulosten analysointi.....	20
4.4.1. Perusarvoinen tarkastelu.....	20
4.4.2. Optimistinen tarkastelu	20
4.4.3. Pessimistinen tarkastelu.....	21
4.4.4. Kokonaisvaltainen tarkastelu	21
5. Yhteenveto ja pohdinta	23
Viitteet	24
Liite A – Esimerkin testausstrategioiden päätöspuu	26
Liite B – Esimerkinomainen hyötyjen laskeminen	27

1. Johdanto

Kliinikot tekevät päivittäin potilaita koskevia päätöksiä arvioituaan ensin potilaan sairauden tilan potilasdiagnoosin avulla. Koska kliinikot eivät aina voi suoraan havainnoida ulkoisesti potilaansa sairauden todellista tilaa, päättelyssä joudutaan usein käyttämään apuna puutteellisia vihjeitä, kuten potilaan historiaa, fyysistä tarkastelua ja diagnostisia testejä. Vaillinaisen informaation perusteella tehtyihin päätöksiin liittyykin aina epävarmuus, jota voidaan kuvata todennäköisyydellä. Potilasdiagnoosin perusteella kliinikko voikin estimoida sairauden todennäköisyyttä kuvaavan prioritodennäköisyyden, jonka perusteella potilaalle valitaan jokin seuraavista testausstrategioista:

1. Potilasta ei hoideta.
2. Hankitaan lisäinformaatiota diagnostisen testin avulla.
3. Aloitetaan potilaan hoito.

Kuhunkin näistä testausstrategioista liittyy odotusarvoinen hyöty, joka kuvaa potilaalle tietyn testausstrategian valinnasta aiheutuvia odotusarvoisia hyötyjä ja haittoja. Odotusarvoinen hyöty riippuu sairauden todellisesta tilasta, jota lääkärin estimoima prioritodennäköisyys kuvaa. Koska kliinikko haluaa tehdä mahdollisimman hyvän päätöksen potilaan kannalta, hän estimoimansa prioritodennäköisyyden avulla maksimoi odotusarvoisen hyödyn valitessaan testausstrategiaa.

Testausstrategian valinta odotusarvoisen hyödyn perusteella ei kuitenkaan ole yksikäsitteistä, sillä sen laskemiseen tarvitaan parametreja, joiden estimoimiseen liittyy epävarmuutta. Esimerkiksi diagnostisiin testeihin liittyvät suureet nettohyöty, tarkkuus sekä erottelukyky, joiden estimointiin liittyy epävarmuutta. Lisäksi tarvitaan tietoa potilaan hoitamisen tai hoitamatta jättämisen hyödyistä sekä terveen että sairaan potilaan tapauksissa. Näitäkään ei voida estimoida yksikäsitteisesti, vaan myös niihin liittyy epävarmuutta.

Tämän työn tavoitteena onkin tutkia, miten edellä kuvattujen parametrien epävarmuuden mallintaminen vaikuttaa testausstrategian valintaan. Työssä tarkastellaan eri testausstrategioiden odotusarvoisia hyötyjä lääkärin estimoiman prioritodennäköisyyden funktiona. Käytännössä tämä toteutetaan luomalla MATLAB-ohjelmistoon käyttöliittymä, joka mahdollistaa epävarmuuksien graafisen tarkastelun.

Kappaleessa 2 luodaan katsaus teoriaan, joka luo pohjan odotusarvoisen hyödyn laskemiseen eri testausstrategioille. Kappaleessa 3 perustellaan ja esitellään parametrien epävarmuuksien mallinnus. Lopuksi kappaleessa 4 tarkastellaan ja analysoidaan epävarmuuden vaikutuksia todelliseen aineistoon perustuvan esimerkin avulla.

2. Teoreettinen tausta

2.1. Prioritodennäköisyys ja sen estimointi

Prioritodennäköisyys $P[D]$ määritellään sairauden D todennäköisyytenä ennen mahdollisen lisäinformaation hankkimista. Diagnostisessa tarkastelussa prioritodennäköisyys toimii siis potilaan sairauden todelliseen tilaan liittyvän epävarmuuden mittana. Prioritodennäköisyys määritellään numerona välillä $[0, 1]$ siten, että

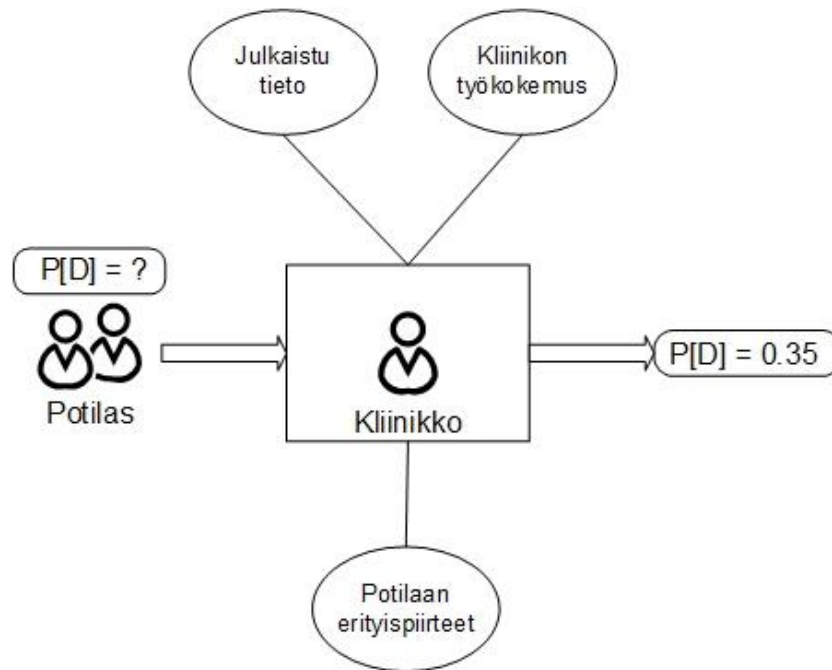
- $P[D] = 1$, jos potilas on varmasti sairas.
- $P[D] = 0$, jos potilas on varmasti ei sairas eli terve.
- Kaikille toistensa poissulkevien tapahtumien summalle $P[D] + P[\text{ei } D] = 1$, missä "ei D " tarkoittaa, että potilas on terve.

Kliinikot estimoivat potilaalle tehtävän diagnoosin yhteydessä sairauden prioritodennäköisyyden, jonka perusteella tehdään päätös potilaalle suoritettavista jatkotoimenpiteistä. Koska potilaan todellinen sairauden tila on usein muiden ulottumattomissa potilaan kehossa, sairautta ei voida aina suoraan havainnoida. Sox, Higgings ja Owens (2013) esittelevät teoksessaan [1] kolme eri menetelmää, jotka auttavat klinikoita prioritodennäköisyyden estimoimisessa. Kuvassa 1 on esitetty estimointia kuvaava prosessi.

Ensimmäinen menetelmä liittyy klinikon omaaman aiemman työkokemuksen hyödyntämiseen. Kokeneet kliinikot pystyvät arvioimaan käsiteltävän potilaan sairauden todennäköisyyttä vertaamalla esimerkiksi potilaan oireita aiempiin samankaltaisiin tapauksiin. Näin saatuun prioritodennäköisyysestimaattiin sisältyy kuitenkin virhettä, sillä klinikolla on käytössään vain rajattu määrä historiallista dataa.

Toinen menetelmä on käyttää hyödyksi yleisesti julkaistua tietoa. Tällöin estimointi voi perustua julkaistuun tietoon esimerkiksi tiettyjen oireiden tai sairauden esiintyvyydestä. Julkaistut raportit ja tutkimukset edustavat myös laajaa kokonaisotosta sairaudesta, minkä ansiosta niiden avulla on myös mahdollista havaita sairauksia, joiden esiintyvyys on pieni. Menetelmään liittyy kuitenkin valintaharha, sillä on mahdollista, että klinikko valitsee tilanteeseen soveltumattomia lähteitä.

Kolmantena menetelmänä on potilaan erityispiirteiden huomioiminen estimoinnissa. Erityispiirteillä tarkoitetaan esimerkiksi potilaan ominaisuuksien tai hänestä helposti saatavien historiatietojen hyödyntämistä. Esimerkiksi potilaan sydän- ja verisuonitautien prioritodennäköisyyden estimoinnissa voidaan ottaa huomioon potilaan aiempi tupakointikäyttäytyminen, painoindeksi, kolesteroli ja verenpaine. Lisäksi sairauden aiempi esiintyvyys potilaan suvussa voi nostaa sairauden prioritodennäköisyyttä, verrattuna mahdollisten tutkimusten antamiin tuloksiin. [1]



Kuva 1: Sairauden D prioritodennäköisyyden $P[D]$ estimointi.

2.2. Posterioritodennäköisyys

Prioritodennäköisyyden perusteella kliinikko voi päättää lisäinformaation hankkimisesta ennen lopullisen hoitopäätöksen tekemistä. Thomas Bayes (1702-61) johti Bayesin teoreeman, jonka avulla voidaan laskea muun muassa sairauden todennäköisyys tämän lisäinformaation hankkimisen jälkeen. [1]

Olkoon A ja B mitä tahansa tapahtumia. Tällöin ehdollinen todennäköisyys $P[A|B]$ määritellään tapahtuman A todennäköisyytenä ehdolla, että tapahtuma B on tosi. Todennäköisysteorian mukaan pätee

$$P[A|B] = \frac{P[A \text{ ja } B]}{P[B]} \quad (1)$$

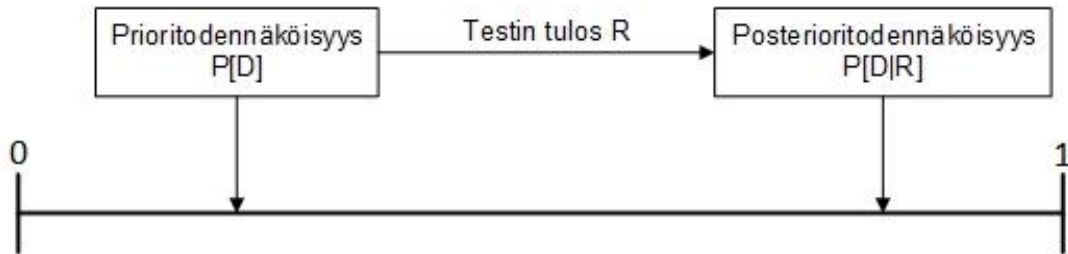
ja

$$P[A] = P[A \text{ ja } B] + P[A \text{ ja } ei B], \quad (2)$$

missä $P[A \text{ ja } ei B]$ on todennäköisyys sille, että tapahtumat A ja $ei B$ tapahtuvat samanaikaisesti, ja $P[B]$ on tapahtuman B todennäköisyys. Tässä $ei B$ tarkoittaa tapahtuman B käänteistapahtumaa eli tapahtumaa, että B ei tapahdu. [2]

Posterioritodennäköisyys $P[D|R]$ määritellään sairauden D todennäköisyytenä lisäinformaation R hankkimisen jälkeen. Posterioritodennäköisyys on siis ehdollinen todennäköisyys, joka voidaan tulkita sairauden D todennäköisyytenä sillä ehdolla, että testin tulos on R . [1] Tässä työssä testin

antama tulos on binäärinen, eli R on joko positiivinen (+), jolloin potilas on testin valossa sairas, tai R on negatiivinen (–), jolloin potilas on testin valossa terve. Kuvassa 2 on esitetty, kuinka prioritodennäköisyys muuttuu posterioritodennäköisyydeksi hankitun lisäinformaation perusteella.



Kuva 2: Sairauden todennäköisyys muuttuu diagnostisessa testissä saadun lisäinformaation ansiosta. Sairauden todennäköisyyttä ennen testin suorittamista kutsutaan prioritodennäköisyydeksi, testin jälkeistä todennäköisyyttä posterioritodennäköisyydeksi.

Merkitään nyt tapahtumia $D =$ ”potilas on sairas”, $ei D =$ ”potilas on terve” ja $R =$ ”testin tulos on R ”. Posterioritodennäköisyyden kertoo tällöin Bayesin teoreema, joka saadaan soveltamalla yhtälöitä (1) ja (2) ehdollisen todennäköisyyden kaavaan. Tällöin posterioritodennäköisyys voidaan lausua muodossa [2]

$$P[D|R] = \frac{P[D \text{ ja } R]}{P[R]} = \frac{P[D \text{ ja } R]}{P[R \text{ ja } D] + P[R \text{ ja } ei D]} = \frac{P[D] \times P[R|D]}{P[D] \times P[R|D] + P[ei D] \times P[R|ei D]}. \quad (3)$$

Posterioritodennäköisyyden laskemiseksi riittää siis seuraavat tiedot [1, 2]:

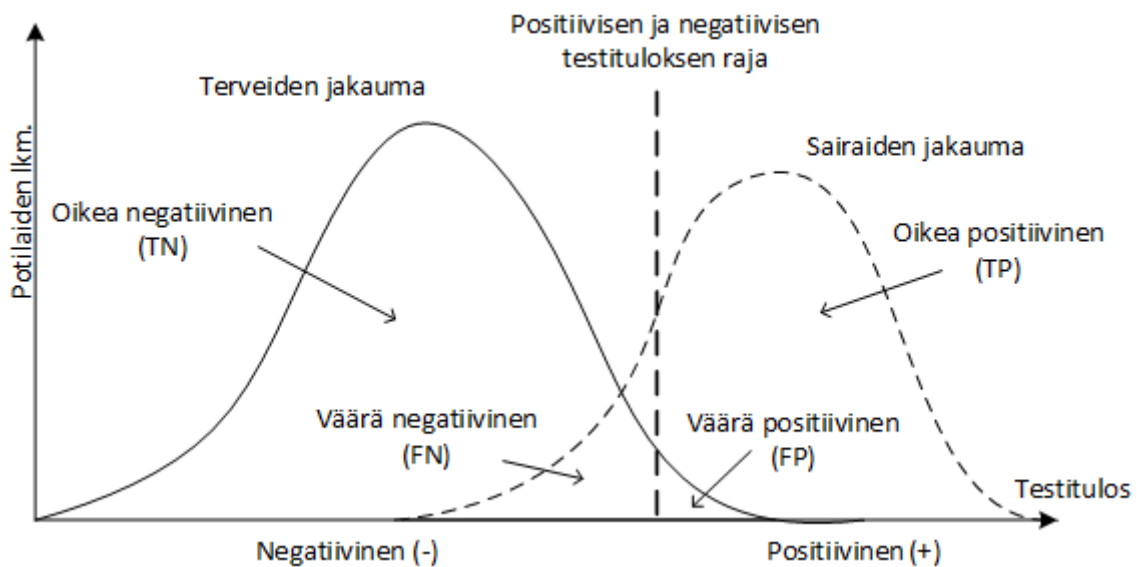
- $P[D]$ eli sairauden prioritodennäköisyys
- $P[R|D]$ eli todennäköisyys, että testin tulos on R (positiivinen tai negatiivinen), kun potilas on sairas
- $P[R|ei D]$ eli todennäköisyys sille, että testin tulos on R (positiivinen tai negatiivinen), kun potilas on terve

Kahden jälkimmäisen todennäköisyyden laskemista käsitellään tarkemmin seuraavassa.

2.3. Testin tarkkuus ja erottelukyky

Simelin ja Rennien (2008) mukaan [3] kliinisessä tarkastelussa diagnostisilla testeillä on kaksi tarkoitusta. Niiden avulla voidaan varmistaa, kärsiikö potilas tietystä sairaudesta, ja toisaalta niiden avulla voidaan varmistaa kyseisen sairauden poissaolo. Yleensä yksi ja sama testi ei sovi molempien

tilanteiden tarkasteluun, sillä sairauden diagnosoiminen vaatii, että positiiviset testitulokset ovat luotettavia, kun taas sairauden poissaolon diagnosoiminen puolestaan vaatii, että testin negatiiviset tulokset ovat luotettavia. [3] Yleensä testin antamat positiiviset ja negatiiviset tulokset eivät ole täysin luotettavia, mitä voidaan havainnollistaa kuvan 3 avulla. Kuvassa on esitetty esimerkinomaisesti terveiden ja sairaiden potilaiden jakauma jatkuvan testituloksen suhteen. Jakaumat ovat usein päällekkäin, jolloin virheellisiltä arvioilta ei voi välttyä. Positiivisen ja negatiivisen testituloksen kynnyksarvoa muuttamalla voidaan vaikuttaa siihen, onko testi luotettavampi sairauden vai sen poissaolon diagnosoimisessa. [1]



Kuva 3: Testituloksen jakaumat sairaalle ja terveelle potilaalle. Sekä positiivinen että negatiivinen testitulos voi olla joko oikea tai väärä.

Päätösmatriisi (taulukko 1) liittyy yhteen diagnostisen testin tuloksen (positiivinen tai negatiivinen) ja sairauden binääriin ulostulon (sairas, terve). Taulukossa 1 on kuvattu kaikki mahdolliset testin jälkeiset tilanteet. Päätösmatriisia voidaan käyttää hyödyksi diagnostisen testin luotettavuutta määrittäessä, kun jokainen taulukon solu edustaa potilaiden absoluuttista lukumäärää kyseisessä tilanteessa. [4]

Taulukko 1: Diagnostisen testin päätöksentekomatriisi

		Potilaan tila		Yhteensä
		Sairas (D)	Terve ($ei D$)	
Testitulos	Positiivinen (+)	a	c	$a + c$
	Negatiivinen (-)	b	d	$b + d$
Yhteensä		$a + b$	$c + d$	$a + b + c + d$

Matriisin avulla voidaan laskea [1, 4] seuraavat neljä lukusuhdetta, jotka kuvaavat myös edellisessä luvussa esitettyjä ehdollisia todennäköisyyksiä $P[R|D]$ tai $P[-R|D]$:

1. Oikeiden positiivisten (True Positive, TP) osuus (TPR) kertoo positiivisten testitulosten osuuden niissä potilaissa, joilla on sairaus. Tätä kutsutaan myös testin tarkkuudeksi (eng. *sensitivity*)

$$TPR = P[+|D] = \text{Tarkkuus} = \frac{a}{a + b}, \quad (4)$$

missä a edustaa sairaita, jotka ovat saaneet positiivisen testituloksen, ja b edustaa terveitä, jotka ovat saaneet positiivisen testituloksen.

2. Väärien positiivisten (False Positive, FP) osuus (FPR) kertoo positiivisten testitulosten osuuden niissä potilaissa, jotka ovat terveitä eli

$$FPR = P[+|ei D] = \frac{c}{c + d}, \quad (5)$$

missä c edustaa terveitä, jotka ovat saaneet positiivisen testituloksen, ja d edustaa terveitä, jotka ovat saaneet negatiivisen testituloksen.

3. Oikeiden negatiivisten (True Negative, TN) osuus (TNR) kertoo negatiivisten testitulosten osuuden niissä potilaissa, jotka ovat terveitä. Tätä kutsutaan myös testin erottelukyvyyksi (eng. *specificity*)

$$TNR = P[-|ei D] = \text{Eroittelukyky} = \frac{d}{c + d}, \quad (6)$$

missä c edustaa terveitä, jotka ovat saaneet positiivisen testituloksen, ja d edustaa terveitä, jotka ovat saaneet negatiivisen testituloksen.

4. Väärien negatiivisten (False Negative, FN) osuus (FNR) kertoo negatiivisten testitulosten osuuden niissä potilaissa, jotka ovat sairaita eli

$$FNR = P[-|D] = \frac{b}{a + b}, \quad (7)$$

missä a edustaa sairaita, jotka ovat saaneet positiivisen testituloksen, ja b edustaa sairaita, jotka ovat saaneet negatiivisen testituloksen.

Yhtälön (3) mukaisesti nyt voidaan laskea [1] posterioritodennäköisyydet sijoittamalla siihen kaavat (4) – (7). Nyt kun R on (+) tai (–), posterioritodennäköisyyksiksi saadaan

$$P[D|+] = \frac{P[D] \times P[+|D]}{P[D] \times P[+|D] + P[ei D] \times P[+|ei D]}, \quad (8)$$

ja

$$P[D|-] = \frac{P[D] \times P[-|D]}{P[D] \times P[-|D] + P[ei D] \times P[-|ei D]} \quad (9)$$

2.4. Odotusarvoinen hyöty ja päätöspuut

Epävarmuuden vallitessa kliininen päätösongelma sisältää seuraavat elementit [5]:

1. Toimenpiteiden eli testausstrategioiden joukko $X = \{x^1 = \text{"Hoida"} (A+), x^2 = \text{"Testaa"} (Test), x^3 = \text{"Älä hoida"} (A-)\}$, joihin päätöksentekijä voi vaikuttaa.
2. Terveystilojen joukko $Y = \{1 = \text{"Sairas"}, 2 = \text{"Terve"}\}$, joihin päätöksentekijä ei voi vaikuttaa.
3. Seurausfunktio $H(x \in X, y \in Y)$, joka kuvaa kaikkia mahdollisia toimenpide-terveydentila-kombinaatioita. Seurausfunktio liittyy jokaiseen näistä testausstrategia-terveydentila-kombinaatioista seurauksen.
4. Todennäköisyysfunktio $P(y \in Y)$, joka kuvaa uskomusta kunkin terveydentilan realisoitumisen todennäköisyydestä. Tilat ovat toisensa poissulkevia, joten pätee $\sum_{y \in Y} P(y) = 1$.
5. Hyötyfunktio $U(H)$, joka mittaa kunkin mahdollisen toimenpide-terveydentila-kombinaation haluttavuutta yksilölle. Hyödyn täytyy olla kardinaalinen eli kvantitatiivisesti mitattavissa.

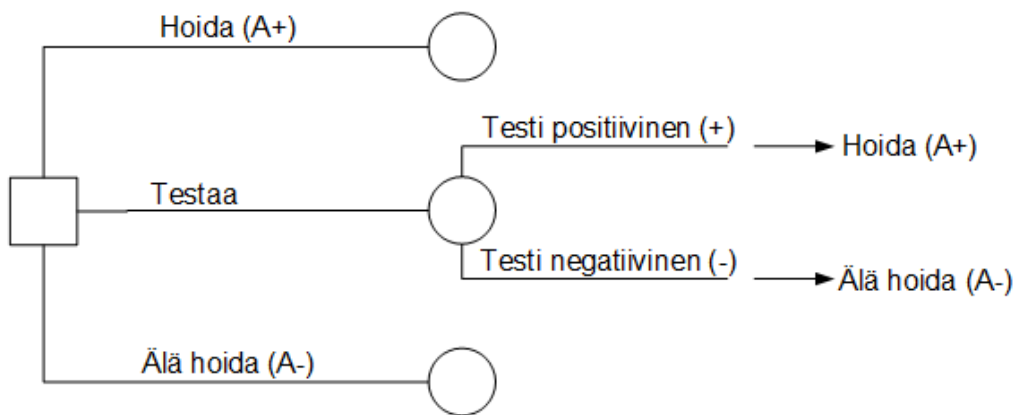
Merkitään kutakin toimenpide-terveydentila-kombinaatiota $H_y^x = (H_1^x, H_2^x)$ ja niitä vastaavia todennäköisyyksiä $p = (p_1, p_2)$. Näiden avulla voidaan nyt määritellä odotusarvoinen hyöty:

$$EU[x] = E[U(H_y^x)] = p_1 U(H_1^x) + p_2 U(H_2^x), \quad (10)$$

missä tapahtuman x odotusarvoinen hyöty $EU[x]$ määritellään siis todennäköisyyksillä painotettuna summana hyötyfunktioista $U(H_y^x)$. [5]

Kliinikon täytyy valita luvussa 2.1. kuvatun prioritodennäköisyyden estimoinnin jälkeen jokin seuraavista potilaalle tehtävistä toimenpiteistä: potilas hoidetaan, potilaasta hankitaan lisätietoa

diagnostisen testin avulla tai potilasta ei hoideta. Jotta toimenpiteen "Testaa" tarkastelu olisi mielekäästä, on testin binäärisille lopputulemille määriteltävä potilaalle suoritettavat jatkotoimenpiteet. Koska tässä työssä on rajattu tarkastelemaan tilannetta, jossa lisättestien suorittaminen ei ole mahdollista, klinikko tekee hoitopäätöksen ensimmäisen testin tuloksen perusteella: mikäli testin tulos on positiivinen (+), potilas hoidetaan (A+), ja mikäli tulos on negatiivinen (-), potilasta ei hoideta (A-). Tilanteesta voidaan piirtää päätöspuu (kuva 4), jossa kukin neliön muotoinen solmu merkitsee päätössolmua, jossa tehdään päätös edellisessä kappaleessa kuvattujen testausstrategioiden X väliltä. Ympyrän muotoinen sattumasolmu kuvaa satunnaisuutta terveydentilojen Y välillä. Kuvien 5 ja 6 päätöspuissa on eritelty tarkemmin kutakin kuvan 4 esittämää testausstrategiaa. Päätöspuiden avulla voidaan kuvata eri testausstrategia-terveydentila-kombinaatiot, ja sitä kautta laskea myös odotusarvoiset hyödyt kullekin testausstrategialle. [1]



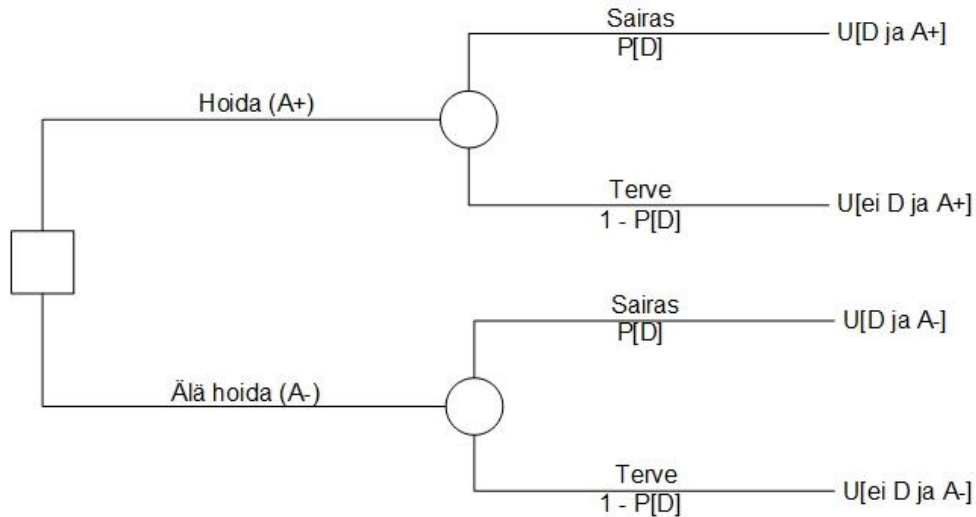
Kuva 4: Kliinikon päätöksentekotilanne, jossa valittavana on kolme eri testausstrategiaa. Strategian "Testaa" tapauksessa positiivinen testitulos indikoi hoidon aloittamista, negatiivinen puolestaan hoitamatta jättämistä.

Tarkastellaan kuvan 5 testausstrategioita "Hoida" (A+) ja "Älä hoida" (A-). Merkitään mahdollisia terveydentiloja "Sairas" (D) ja "Terve" (ei D). Lisäksi määritellään hoitamisen hyödyt ($U[D \text{ ja } A+]$ ja $U[ei D \text{ ja } A+]$) sekä hoitamatta jättämisen hyödyt ($U[D \text{ ja } A-]$ ja $U[ei D \text{ ja } A-]$), jotka kuvaavat eri toimenpide-terveydentila-kombinaatioiden yksilöllisiä ja yhteiskunnallisia seurauksia. Nämä hyödyt määritellään välillä $[0,1]$, missä 0 = "äärimmäinen haitta (kuolema)" ja 1 = "äärimmäinen hyöty (täydellinen terveys ilman haittoja)". Esimerkiksi $U[D \text{ ja } A-]$ kuvaa sairaan henkilön hoitamatta jättämisestä aiheutuvia seurauksia, jolloin hyötyjen laskennassa voidaan ottaa huomioon muun muassa yksilön odotusarvoinen elinikä sekä yhteiskunnalle aiheutuvat rahalliset kustannukset. Tapahtuman "Hoida" odotusarvoinen hyöty voidaan nyt esittää yhtälön (10) mukaisesti

$$\begin{aligned} EU[A+] &= P[D] \times U[D \text{ ja } A+] + P[ei D] \times U[ei D \text{ ja } A+] \\ &= P[D] \times U[D \text{ ja } A+] + (1 - P[D]) \times U[ei D \text{ ja } A+], \end{aligned} \quad (11)$$

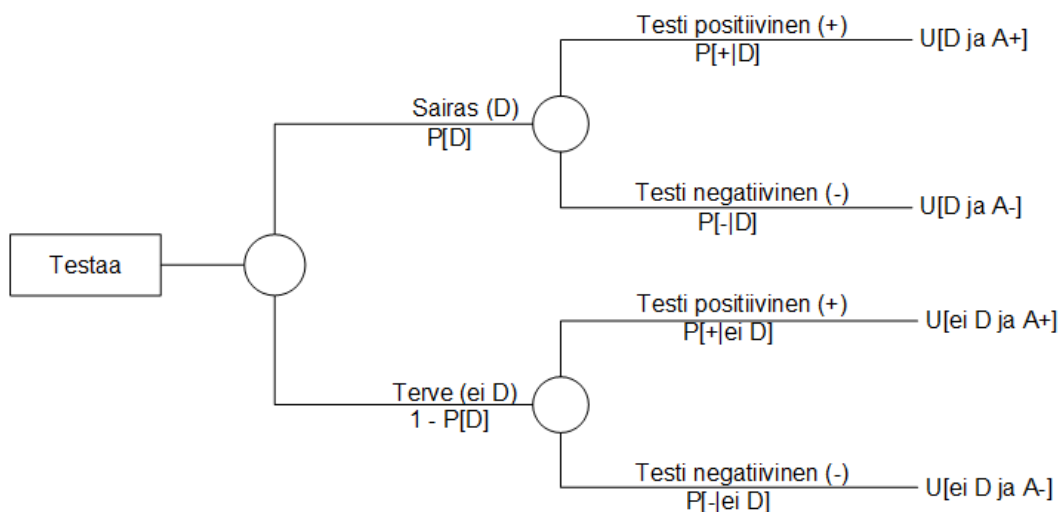
ja vastaavasti tapahtumalle "Älä hoida"

$$EU[A -] = P[D] \times U(D \text{ ja } A -) + (1 - P[D]) \times U(\text{ei } D \text{ ja } A -). \quad [1] \quad (12)$$



Kuva 5: Päätöstilanne, jossa klinikolla on valintavaihtoehtoina joko hoitaa tai olla hoitamatta potilasta. Molemmissa tilanteissa potilas voi olla joko sairas tai terve.

Jos toimenpiteenä on testaus, tilanne on hieman monimutkaisempi. Testauspäätökseen liittyy epävarmuutta terveydentilan suhteen, sillä potilas on joko "Sairas" (D) tai "Terve" ($\text{ei } D$) niitä vastaavilla todennäköisyyksillä $P[D]$ ja $1 - P[D]$. Tämän lisäksi testin tulos voi olla molempien terveydentilojen suhteen joko positiivinen (+) (potilas hoidetaan) tai negatiivinen (-) (potilasta ei hoideta). Yhteensä saadaan siis neljä eri kombinaatiota, jotka on esitetty kuvassa 6.



Kuva 6: Päätöspuu, kun klinikko on valinnut testausstrategiaksi lisäinformaation hankkimisen diagnostisen testin avulla.

Strategian ”Testaa” kohdalla testin antaessa positiivinen tuloksen (riippumatta onko henkilö sairas vai terve), suoritetaan potilaalle hoito, jolloin hyöty on sairaan potilaan tapauksessa $U[D \text{ ja } A+]$ ja terveen potilaan tapauksessa $U[ei D \text{ ja } A+]$. Vastaavasti negatiivisen testituloksen tapauksessa potilas jätetään hoitamatta, jolloin hyödyt ovat $U[D \text{ ja } A-]$ ja $U[ei D \text{ ja } A-]$. [1]

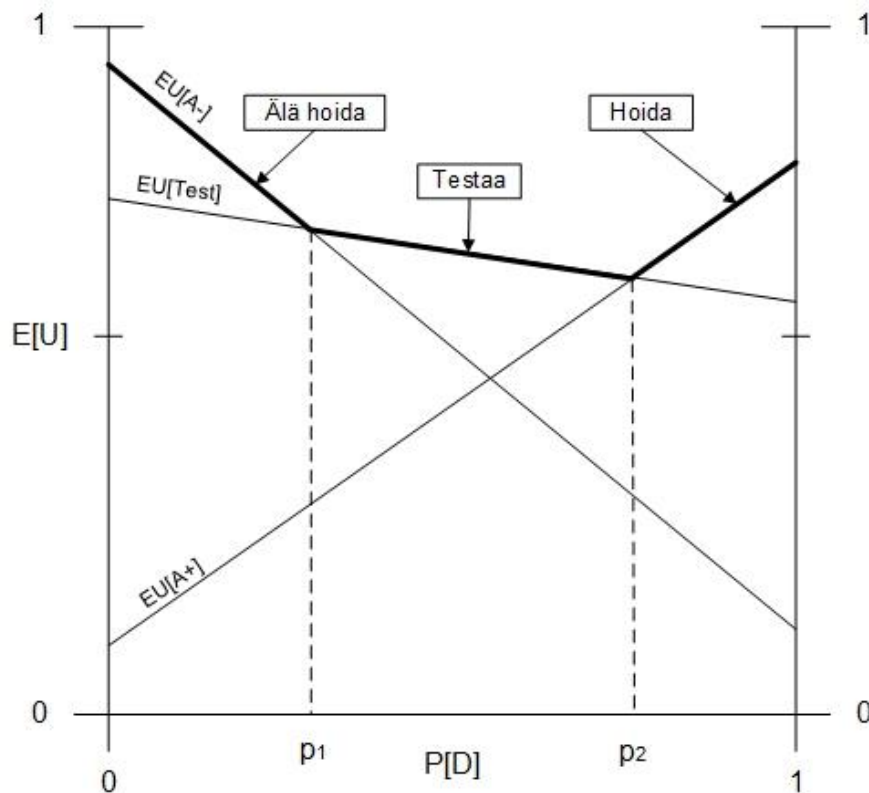
Testin itsensä suorittamiseen liittyy myös sairaudesta irrallinen nettohyöty $U[T] \in [-1,1]$, joka määritellään testin suorittamisesta aiheutuvina yksilöllisinä ja yhteiskunnallisina seurauksina [1, 6]. Testin suorittaminen voi esimerkiksi aiheuttaa potilaalle infektio- tai kuolemanriskin, jolloin testillä on suora vaikutus potilaan terveydentilaan. Testillä voi olla myös välillisiä vaikutuksia, sillä esimerkiksi testitulosten odotus voi viivästyttää potilaan hoidon aloittamista, ja lisäksi odotus voi aiheuttaa potilaalle psykologista harmia. Testillä voi myös olla taloudellisia kustannuksia sekä yksilön että yhteiskunnan näkökulmasta. Testin nettohyöty on summa testin suorittamisesta aiheutuvista positiivisista vaikutuksista (esimerkiksi epä tietoisuudessa elävä potilas saa rauhan) ja negatiivisista vaikutuksista (esimerkiksi testin aiheuttama fyysinen haitta potilaalle). Jos nettohyöty on negatiivinen, niin nettovaikutus testin kokemisesta ei ole haluttava, ja positiiviselle arvolle päinvastoin. Testausstrategian ”Testaa” (*Test*) odotusarvoinen hyöty voidaan esittää

$$\begin{aligned}
 EU[Test] = & P[D] \times (P[+|D] \times U[D \text{ ja } A+] + P[-|D] \times U[D \text{ ja } A-]) \\
 & + (1 - P[D]) \\
 & \times (P[+|ei D] \times U[ei D \text{ ja } A+] + P[-|ei D] \times U[ei D \text{ ja } A-]) \\
 & + U[T]. \quad [1]
 \end{aligned} \tag{13}$$

Kaavoista (11), (12) ja (13) huomataan, että kunkin strategian odotusarvoinen hyöty voidaan kuvata prioritodennäköisyyden $P[D]$ lineaarisena funktiona, kun tiedetään tarvittavat sairauksiin ja testeihin liittyvät hyödyt. Piirtämällä strategiat prioritodennäköisyyden funktiona, saadaan tuloksena kuvan 7 esittämät kolme suoraa, jotka ovat lineaarisia prioritodennäköisyyden suhteen.

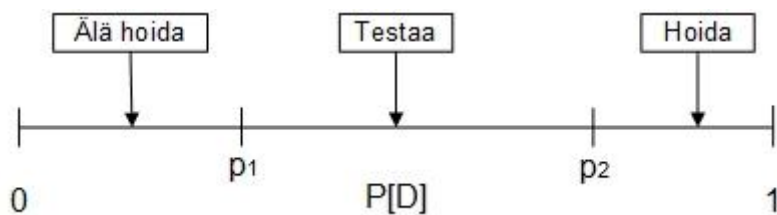
Koska testausstrategioiden odotusarvoiset hyödyt mittaavat kunkin mahdollisen strategian odotusarvoista haluttavuutta potilaalle, potilaan kannalta paras strategia on se, joka maksimoi odotusarvoisen hyödyn kullakin prioritodennäköisyyden arvolla [1]. Kuvasta 7 huomataan, että suurin odotusarvoinen hyöty prioritodennäköisyyden välillä $[0, p_1]$ on toimenpiteellä ”Älä hoida”, välillä $[p_1, p_2]$ toimenpiteellä ”Testaa” ja välillä $[p_2, 1]$ toimenpiteellä ”Hoida”. Maksimoitua odotusarvoista hyötyä on esitetty kuvassa 7 tummennetulla viivalla.

Jos testin nettohyöty $U[T] = 0$, niin prioritodennäköisyydet p_1 ja p_2 ovat rajapisteitä sille, että testi voi siirtää priorin toiselle puolelle hoidon kynnyksarvoa (kuvan 7 suorien $EU[A-]$ ja $EU[A+]$ leikkauspisteen $P[D]$ -koordinaatti). Siten p_1 on pienin priorin, josta positiivisen testituloksen kautta posterioritodennäköisyys päivittyy suuremmaksi kuin hoidon kynnyksarvo. Vastaavasti p_2 on suurin priorin, jolla negatiivisen testituloksen kautta posterior päivittyy pienemmäksi kuin hoidon kynnyksarvo. Näin ollen strategian ”Testaa” suorittamisesta ei ole hyötyä väleillä $[0, p_1]$ ja $[p_2, 1]$.



Kuva 7: Strategioita "Älä hoida", "Testaa" ja "Hoida" vastaavat odotusarvoiset hyödyt $EU[A-]$, $EU[Test]$ ja $EU[A+]$ prioritodennäköisyyden $P[D]$ funktiona.

Edellä esitetyn perusteella kliinikon testausstrategian valintaa voidaan kuvata janan, jossa potilaan kannalta optimaalinen testausstrategia riippuu prioritodennäköisyyden väleistä. Kuvaa 7 vastaava optimaalinen jana on esitetty alla olevassa kuvassa 8.



Kuva 8: Potilaan odotusarvoisen hyödyn maksimoiva testausstrategia kliinikon estimoiman prioritodennäköisyyden funktiona.

3. Epävarmuuden huomioiminen

3.1. Perustelut epävarmuuden mallintamiselle

Diagnostisten testien tarkkuutta TPR (kaava (4)) ja erottelukykyä FNR (kaava (6)) mitataan alun perin vertailemalla tutkittavan diagnostisen testin tuloksia ”sairas” ja ”terve” -testiryhmien välillä. Molemmat ryhmät käyvät läpi myös ”kultaisen standardin” (eng. *gold standard*) testin, joka toimii tutkittavan testin vertailupohjana. Koska vertailupohjana käytettyyn kultaisen standardin testin tarkkuuteen ja erottelukykyyhin liittyy epävarmuutta, myös tutkittavan testin tuloksia kuvaavan taulukon 1 jokaiseen neljään soluun (luvut a, b, c, d) liittyy epävarmuutta. Testin tarkkuuteen ja erottelukykyyhin tuo epävarmuutta myös käytetty testiryhmäotos. [1, 7]

Diagnostisen testin tarkkuus ja erottelukyky voidaan estimoida myös käyttämällä hyväksi jo julkaistua tutkimustietoa. Samasta testistä saattaa kuitenkin olla tehty useita eri tutkimuksia, joiden antamat testien tarkkuudet ja erottelukyvyyt eroavat toisistaan. Tutkimustuloksia voidaan tällöin yhdistellä, eli suorittaa niin sanottu meta-analyysi, jonka tekemiseen liittyy yhdistelyn vuoksi olennaisesti epävarmuus. [7]

Testin nettohyöty $U[T]$ kertoo, mitkä ovat testin suorittamisesta aiheutuvat yksilölliset ja yhteiskunnalliset seuraukset. Yksilön kannalta nettohyödyssä tulee arvioida esimerkiksi testin emotionaaliset, sosiaaliset, kognitiiviset, käyttäytymiselliset ja terveydelliset vaikutukset. Toisaalta testistä voi myös aiheutua yhteiskunnallisia kustannuksia, jotka tulee ottaa huomioon. Testin nettohyödyn määrittämiseen voidaan käyttää esimerkiksi satunnaistettuja kokeita tai erilaisia päätöksentekomalleja, mutta hyödyn määrittäminen ei kuitenkaan ole yksikäsitteistä, sillä parhaimmillaankin testin nettohyöty kuvaa vain testin odotusarvoista hyötyä tietylle populaatiolle. Todellisuudessa testin seuraukset ovat kuitenkin yksilö- ja populaatioriippuvaisia, ja siten ollen myös testin nettohyödyn estimointiin liittyy epävarmuutta. [6]

Hoitamisen ($U[D \text{ ja } A+]$ ja $U[ei D \text{ ja } A+]$) sekä hoitamatta jättämisen ($U[D \text{ ja } A-]$ ja $U[ei D \text{ ja } A-]$) odotusarvoiset hyödyt eri terveydentiloille estimoidaan tutkimalla esimerkiksi sairauden aiheuttamia pidemmän ja lyhyemmän aikavälin seurauksia sekä yhteiskunnalle että potilaalle. Estimointiin voidaan tehdä esimerkiksi subjektiivisesti trade-off -menetelmien avulla tai eliniänodotteiden avulla laskemalla. [1] Kummassakaan tapauksessa hyötyjen määrittäminen ei kuitenkaan ole yksikäsitteistä, minkä vuoksi epävarmuuden mallintaminen myös näihin parametreihin on perusteltua.

3.2. Parametrien mallinnus

Kaikki tässä työssä tarkasteltavat parametrit on esitelty jo aiemmin teoriaosuuden eri kappaleissa. Näiden parametrien mahdolliset vaihteluvälit ja tulkinnat on puolestaan esitetty taulukossa 2 [1].

Taulukko 2: Epävarmuusparametrien mahdolliset vaihteluvälit ja tulkinnat.

Parametri	Vaihteluväli	Parametrin tulkinta
a, b, c ja d	$a > b > 0$ $d > c > 0$	Kombinaatioiden $(+, D)$, $(-, D)$, $(+, ei D)$ ja $(-, ei D)$ otosmäärät testin tarkkuutta ja erottelukykyä määritettäessä. Testin tarkoituksenmukaisuuden vuoksi $a > b$ ja $d > c$ [3].
$U[D$ ja $A+$]	$[0, 1]$	Sairaahan henkilön hoitamisesta aiheutuvat yksilölliset ja yhteiskunnalliset seuraukset.
$U[ei D$ ja $A+$]	$[0, 1]$	Terveen henkilön hoitamisesta aiheutuvat yksilölliset ja yhteiskunnalliset seuraukset.
$U[D$ ja $A-$]	$[0, 1]$	Sairaahan henkilön hoitamatta jättämisestä aiheutuvat yksilölliset ja yhteiskunnalliset seuraukset.
$U[ei D$ ja $A-$]	$[0, 1]$	Terveen henkilön hoitamatta jättämisestä aiheutuvat yksilölliset ja yhteiskunnalliset seuraukset.
$U[T]$	$[-1, 1]$	Testin suorittamisesta aiheutuvat yksilölliset ja yhteiskunnalliset seuraukset.

Tässä työssä edellä esitettyjen parametrien epävarmuutta on mallinnettu väleinä. Ensin parametrille on annettu perusarvo, jonka jälkeen tämän arvon on annettu liikkua tietyllä prosenttivälillä sekä ala- että yläpuolelta. Esimerkiksi testin nettohyödyille voidaan antaa ensin perusarvo $U_B[T] = -0.5$, minkä jälkeen sen annetaan liikkua siten, että $U[T] \in [1.10 \times U_B[T], 0.95 \times U_B[T]] = [U_{min}[T], U_{max}[T]] = [-0.550, -0.475]$ on intervalli. Koska kaikkien yhdeksän epävarman parametrin annetaan liikkua samanaikaisesti tietyillä intervaleilla, on eri strategioiden odotusarvoisten hyötyfunktioiden (kaavat (11) – (13)) rajattu lähtöjoukko

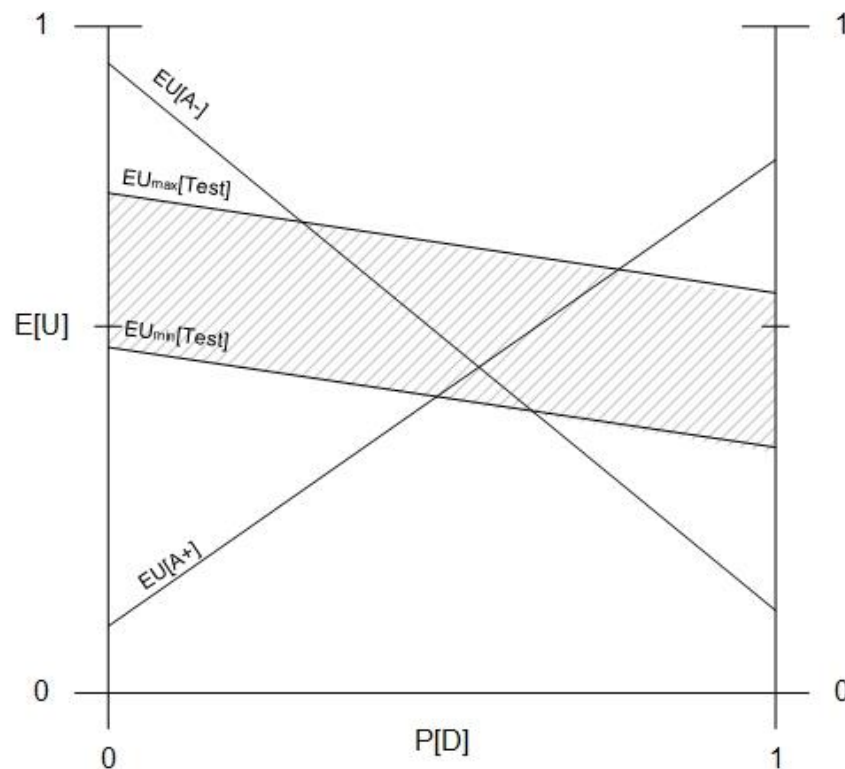
$$S = (a, b, c, d, U[T], U[D \text{ ja } A+], U[ei D \text{ ja } A+], U[D \text{ ja } A-], U[ei D \text{ ja } A-]).$$

Merkitään testausstrategioiden joukosta $X = \{x^1 = \text{"Hoida"} (A+), x^2 = \text{"Testaa"} (Test), x^3 = \text{"Älä hoida"} (A-)\}$ valittua yksittäistä testausstrategiaa $x^i \in X$, missä $i = 1, 2, 3$. Nyt koska parametreja mallinnetaan intervaleilla, myös kunkin testausstrategiavaihtoehdon x^i odotusarvoista hyötyä $EU[x^i, P[D]]$ voidaan kuvata intervaleilla $P[D]$:n funktiona. Lineaariset funktiot saavuttavat maksiminsa ja miniminsä joukon S rajapisteissä [8], joten minimi ja maksimi lineaariselle funktioille $EU[x^i, P[D]]$ saadaan tutkimalla kunkin parametrin yksittäistä vaikutusta eri strategioiden odotusarvoisiin hyötyihin kaavojen (11) – (13) mukaisesti (taulukko 3). Näin voidaan valita kustakin parametrilla aina joko ala- tai yläraja, jolloin kaikille $i = 1, 2, 3$ saadaan laskettua $EU_{min}[x^i, P[D]] = \min_S \{EU[x^i, P[D]]\}$ ja $EU_{max}[x^i, P[D]] = \max_S \{EU[x^i, P[D]]\}$ valitsemalla tarpeen mukaan hyödyn minimoiva tai maksimoiva parametrin arvo. Esimerkiksi testauksen ($x^2 = Test$) odotusarvoinen hyöty $EU[Test, P[D]]$ maksimoituu parametrin b ja $U[T]$ suhteen, kun valitaan aina parametrin b alaraja b_{min} ja nettohyödyn $U[T]$ yläraja $U_{max}[T]$.

Taulukko 3: Parametrin kasvun vaikutus odotusarvoisiin hyötyihin.

Parametrin kasvu	Vaikutus odotusarvoiseen hyötyyn		
	$EU[A+]$	$EU[A-]$	$EU[Test]$
a	ei vaikutusta	ei vaikutusta	+
b	ei vaikutusta	ei vaikutusta	-
c	ei vaikutusta	ei vaikutusta	-
d	ei vaikutusta	ei vaikutusta	+
$U[T]$	ei vaikutusta	ei vaikutusta	+
$U[D \text{ ja } A+]$	+	ei vaikutusta	+
$U[\text{ei } D \text{ ja } A+]$	+	ei vaikutusta	+
$U[D \text{ ja } A-]$	ei vaikutusta	+	+
$U[\text{ei } D \text{ ja } A-]$	ei vaikutusta	+	+

Odotusarvoinen hyöty kaikille testausstrategioille $x^i \in X$ voidaan siis esittää kullakin $P[D]$:n arvolla välinä $EU[x^i, P[D]] \in [EU_{min}[x^i, P[D]], EU_{max}[x^i, P[D]]]$, missä minimi ja maksimi joukossa S saadaan laskettua aiemmassa kappaleessa kuvatulla tavalla. Kuvassa 9 on esitetty graafisesti pelkästään testin nettohyödyn $U[T]$ epävarmuuden vaikutus testin odotusarvoiseen hyötyyn $EU[Test]$, joka on merkitty kuvaan tummennettuna alueena. Testin odotusarvoisen hyödyn riippuvuus nettohyödyistä ilmenee myös kaavasta (13).



Kuva 9: Nettohyödyn $U[T]$ epävarmuuden vaikutus testin odotusarvoiseen hyötyyn $EU[Test]$. Suora $EU_{max}[Test]$ kuvaa maksimiarvoa $P[D]$:n funktiona, $EU_{min}[Test]$ puolestaan minimiarvoa $P[D]$:n funktiona.

3.3. Dominanssi

Dominanssi määritellään relaationa, jonka avulla eri testausstrategiavaihtoehtoja voidaan verrata toisiinsa epävarmuuden vallitessa [8]. Nyt strategia x^i dominoi strategiaa x^j välillä $P[D] \in [q, w]$, jos seuraavat kaksi ehtoa pätevät [8, 9]:

$$1) EU_{min}[x^i, P[D]] \geq EU_{max}[x^j, P[D]] \text{ kaikilla } P[D] \in [q, w].$$

$$2) EU_{min}[x^i, P[D]] > EU_{max}[x^j, P[D]] \text{ jollakin } P[D] \in [q, w].$$

Jos nyt nämä kaksi ehtoa pätevät siten, että jokin strategia x^i dominoi kaikkia muita strategioita $x^j, j = 1, 2, \dots, n \neq i$ välillä $[q, w]$, on x^i tällöin aina parempi kuin muut strategiat kyseisellä välillä [9]. Näin ollen dominanssien avulla on mahdollista löytää prioritodennäköisyyden $P[D] \in [q, w]$ väleille testausstrategioita, jotka ovat aina kyseisellä välillä parempia kuin muut strategiat [8].

4. Esimerkki

4.1. Taustaa

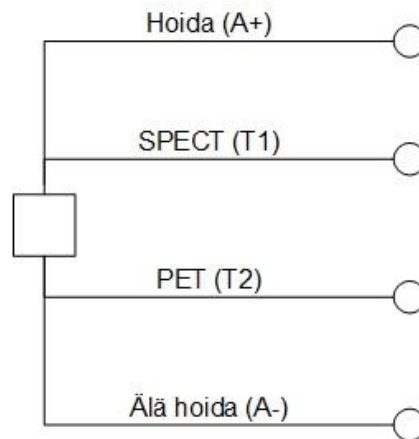
Tässä työssä on tutkittu epävarmuuden huomioimista testausstrategian valinnassa, kun klinikon valittavissa on hoidon tai hoitamatta jättämisen lisäksi kaksi eri diagnostista testiä lisäinformaation hankkimista varten. Esimerkissä on käytetty hyödyksi Amerikkalaisen sydänjärjestön (American Heart Association, 2012) julkaisemaa meta-analyysiä PET (Positron Emission Tomography) ja SPECT (Single Photon Emission Computed Tomography) –testeistä, joilla molemmilla voidaan tutkia sepelvaltimotautia [10]. Huomioitavaa on, että esimerkissä käytettävä aineisto ei suoraan anna kaikkia tarvittavia tietoja, vaan esimerkiksi testien ja sairauksien hyödyt on laskettu esimerkinomaisesti käyttämällä hyödyksi yleistä tietoa sepelvaltimotaudista (lähteet [11] – [14], laskut liitteessä B). Käytännössä epävarmuuksia on tässä esimerkissä tutkittu käyttämällä hyödyksi Matlab-ohjelmistoon luotua graafista käyttöliittymää.

4.2. Rajaukset

Tähän mennessä tässä työssä on tehty testausstrategian valintaan liittyen useita rajoituksia, jotka on syytä kerrata ennen tarkempaa syventymistä esimerkkiin. Ensinnäkin potilaan sairauden tilan on oletettu olevan binäärinen, eli potilas voi olla ainoastaan sairas tai terve. Toisekseen, myös testien oletetaan olevan binäärisiä, eli testin tulos voi olla ainoastaan positiivinen tai negatiivinen. Lisäksi tässä työssä on rajoitettu testausstrategian tarkastelu vain yhteen vaiheeseen, jossa kahden eri testin samanaikainen tai peräkkäinen suorittaminen ei ole mahdollista. Tässä työssä oletetaan myös sairauden esiintyvyyden olevan yhtä suurta klinikon vastaanotolle tulevassa populaatiossa kuin testin tunnuslukujen mittaamisessa käytetyssä populaatiossa, mikä oikeuttaa testin tarkkuuden ja erottelevuuden käyttöön jatkotarkasteluissa [4].

4.3. Testiasetelma ja tulokset

Tässä esimerkissä klinikko epäilee potilaalle kuvan 1 mukaan tehdyn diagnostisen tarkastelun perusteella, että potilaalla saattaisi olla sepelvaltimotauti (eng. CAD, Coronary Arterary Disease) todennäköisyydellä $P[D]$. Kliinikko voi päättää hoidon aloittamisesta, hoitamatta jättämisestä, PET –testiin lähettämisestä tai SPECT –testiin lähettämisestä. Kliinikko valitsee siis jonkin kuvan 10 testausstrategioista. Kokonaisuudessaan testausstrategia on esitetty liitteen A päätöspuussa, josta ilmenee todennäköisyyksien lisäksi myös kaikki mahdolliset lopputulemat.



Kuva 10: Vaihtoehtoiset testausstrategiat esimerkin tapauksessa. Potilas voidaan hoitaa, lähettää jompaan kumpaan testiin tai olla hoitamatta.

Luvussa 3 esitettyjen syiden perusteella testien tarkkuuteen ja erottelukykyyyn liittyy epävarmuuksia. Tässä esimerkissä testien tarkkuudet ja erottelukyvyyt on saatu Amerikkalaisen Sydänjärjestön meta-analyysistä [10], jossa on yhdistelty molempien testien tapauksessa useamman eri tutkimusten tuloksia. Luottamusvälit (95 %) molemmille testien tarkkuuksille ja erottelukyvyyille on meta-analyysin mukaan saatu tutkimalla testituloksien variansseja ns. *logit* –funktion avulla. Testien tarkkuuteen ja erottelukykyyyn liittyvä data on esitetty taulukossa 4.

Taulukko 4: Esimerkissä käytettyjen kahden eri testin (SPECT ja PET) tarkkuudet sekä erottelukyvyyt 95 %:n luottamusväleillä.

n = yhdistelyjen tutkimusten lkm.	Tarkkuus (%) $P[+ D]$ (Alaraja, Yläaraja)	Erottelukyky (%) $P[- e\bar{i} D]$ (Alaraja, Yläaraja)
SPECT ($n = 115$)	88.3 (86.5, 90.0)	76.0 (72.4, 79.4)
PET ($n = 9$)	92.6 (88.3, 95.5)	81.3 (66.6, 90.4)

Koska tarkkuus ja erottelukyky riippuvat kaavojen (4) ja (6) mukaisesti suhteista $\frac{a}{a+b}$ sekä $\frac{d}{c+d}$, riittää tarkastella näiden suhteiden välejä eli voidaan valita parametreille b ja c mielivaltaiset arvot, ja

tämän jälkeen antaa välit parametreille a ja d . Välit määritellään siten, että tarkkuuden ja erottelukyvyn ala- ja ylärajat säilyvät taulukon 4 mukaisina. Taulukossa 5 on esitetty parametreille annetut arvot sekä välit.

Taulukko 5: Esimerkissä käytettyjen parametrien välit, kun testien tarkkuudet ja erottelukyvyt ovat taulukon 4 mukaisilla väleillä.

Testi \ Parametri	a Perusarvo (Alaraja, Ylärajaja)	b	c	d Perusarvo (Alaraja, Ylärajaja)
SPECT ($n = 115$)	755 (641, 900)	100	100	317 (262, 385)
PET ($n = 9$)	1250 (755, 2120)	100	100	435 (199, 942)

Koska tarkkoja estimaatteja testien nettohyödyistä $U[PET]$ ja $U[SPECT]$ ei ole saatavilla, oletetaan ne annetuiksi taulukossa 6. Liitteessä B on esimerkinomaisesti laskettu ja esitetty, miten taulukossa annettuihin tuloksiin voitaisiin päästä, ja minkälaisia asioita hyödyn laskemisessa tulisi ottaa huomioon.

Taulukko 6: Testien nettohyödyt $U[PET]$ ja $U[SPECT]$ ala- ja ylärajoineen.

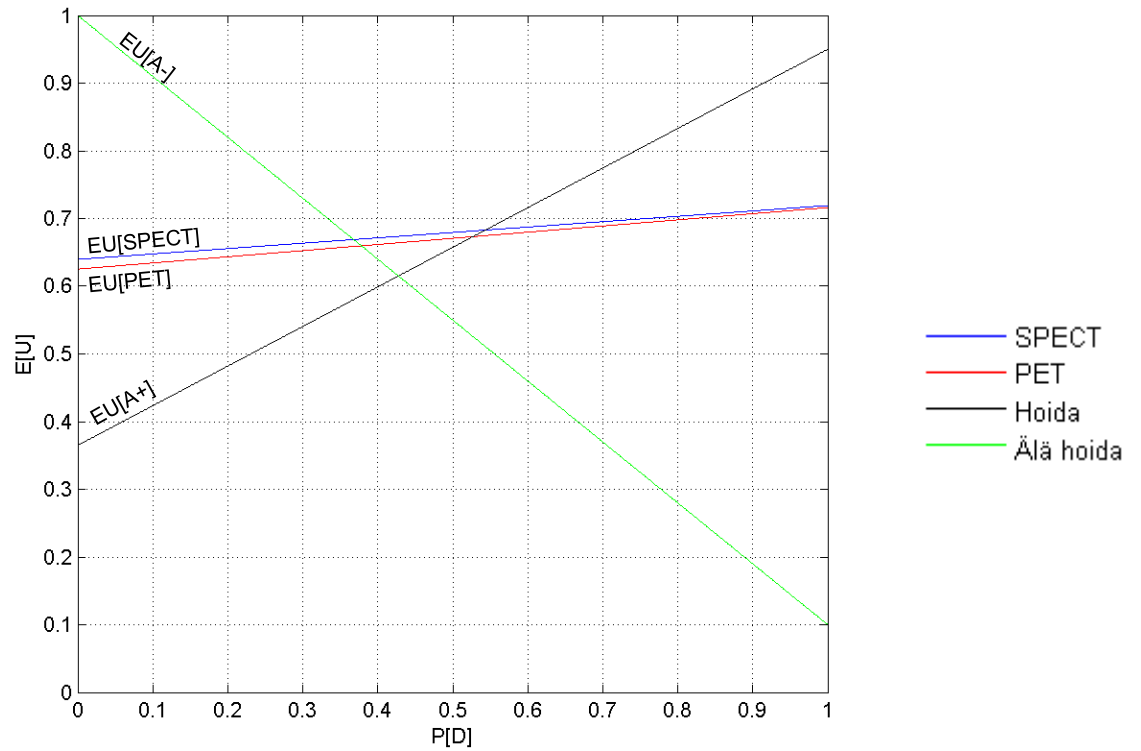
Testi \ Parametri	$U[PET]$	$U[SPECT]$
Nettohyöty (Alaraja, Ylärajaja)	-0.064 (-0.067, -0.061)	-0.052 (-0.055, -0.049)

Oletetaan myös toimenpide-terveydentila-kombinaatioiden hyödyt $U[D ja A+]$, $U[ei D ja A+]$, $U[D ja A-]$ ja $U[ei D ja A-]$ annetuiksi taulukossa 7. Myös näille hyödyille on esitetty esimerkinomainen laskenta liitteessä B.

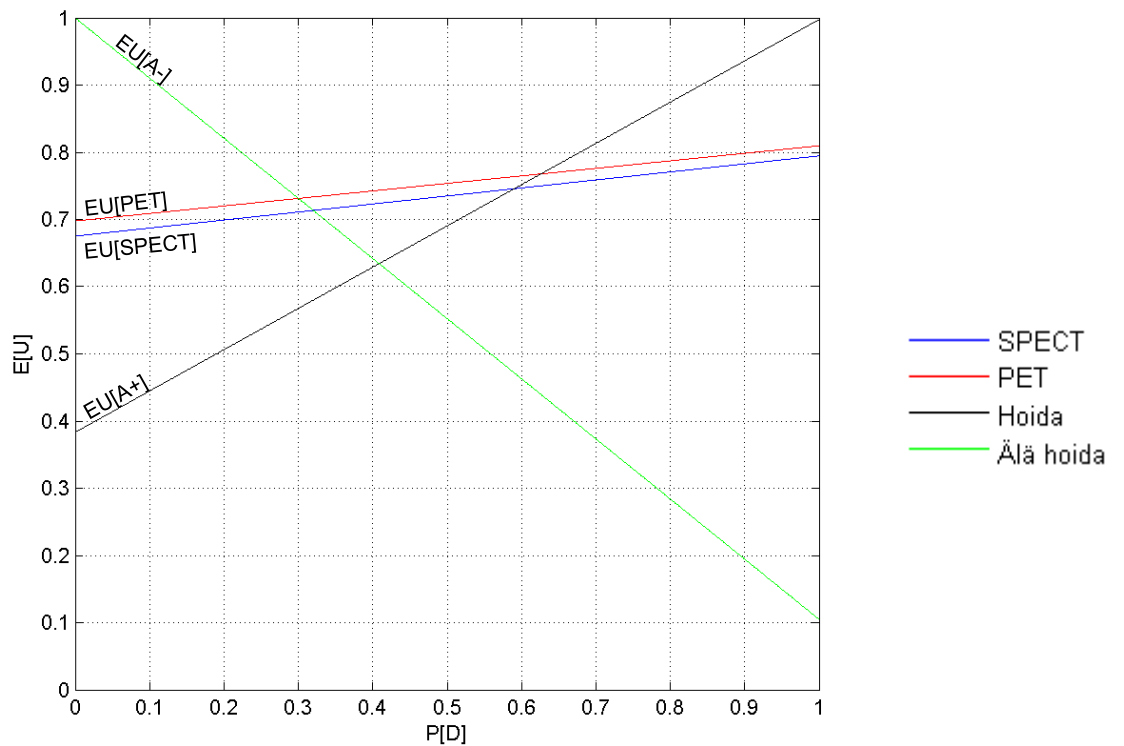
Taulukko 7: Esimerkissä käytettyjen parametrien välit, kun testien tarkkuudet ja erottelukyvyt ovat taulukon 4 mukaisilla väleillä.

	$U[D ja A+]$	$U[ei D ja A+]$	$U[D ja A-]$	$U[ei D ja A-]$
Hyöty (Alaraja, Ylärajaja)	0.95 (0.90, 1.00)	0.37 (0.35, 0.39)	0.10 (0.095, 0.11)	1 (0.95, 1.00)

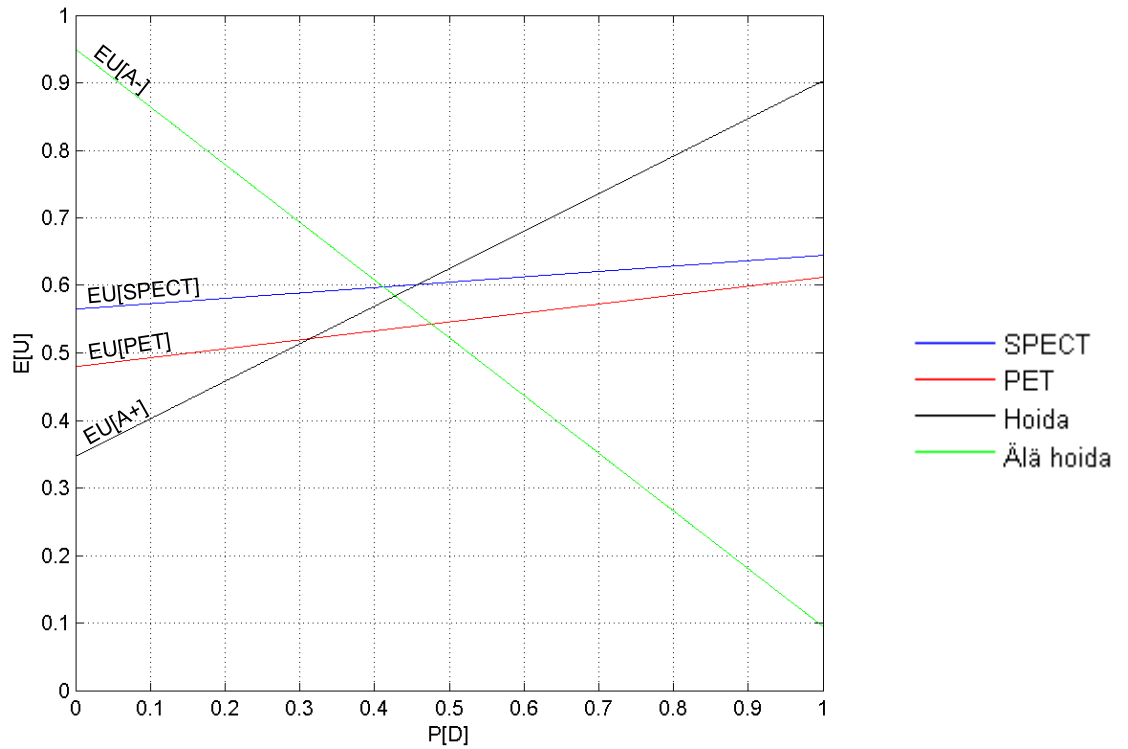
Käyttämällä edellä esitettyjä taulukoiden 5-7 perusarvoja, kunkin testausstrategian odotusarvoiset hyödyt voidaan nyt laskea $P[D]$:n funktiona, ja esittää ne kuvassa 11. Hyödyntämällä taulukkoa 3 voidaan laskea ja esittää myös eri strategioiden odotusarvoisten hyötyjen maksimit (kuva 12) ja minimi (kuva 13) kullakin $P[D]$:n arvolla. Kokonaisuudessaan testausstrategioiden odotusarvoiset hyödyt vaihteluväleineen on esitetty kuvassa 14.



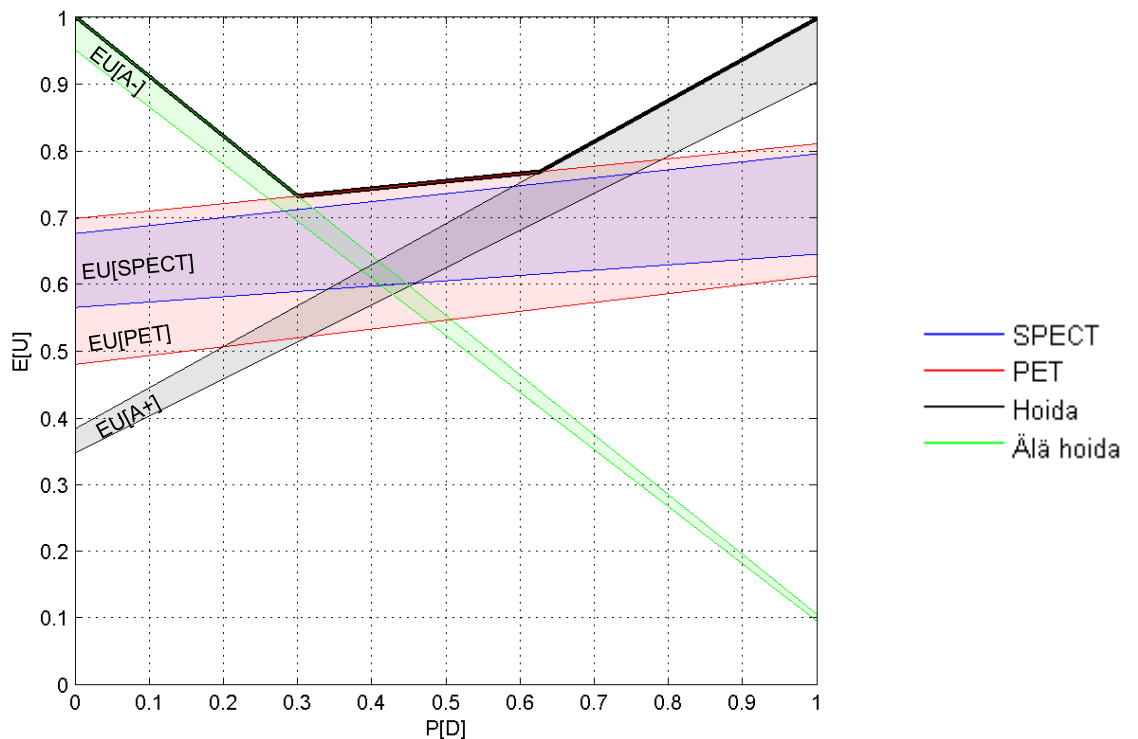
Kuva 11: Vaihtoehtoisten testausstrategioiden odotusarvoisten hyötyjen perusarvot prioritodennäköisyyden funktiona.



Kuva 12: Vaihtoehtoisten testausstrategioiden odotusarvoisten hyötyjen maksimit prioritodennäköisyyden funktiona.



Kuva 13: Vaihtoehtoisten testausstrategioiden odotusarvoisten hyötyjen minimit prioritodennäköisyyden funktiona.

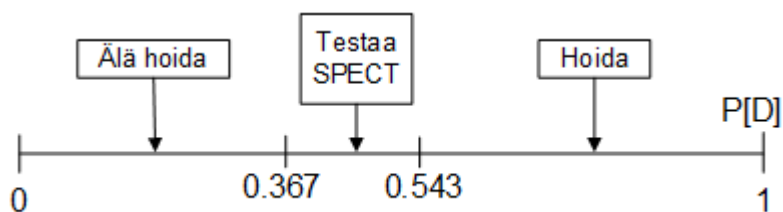


Kuva 14: Vaihtoehtoisten testausstrategioiden odotusarvoiset hyödyt väleineen. Odotusarvoisten hyötyjen maksimiarvot kullakin $P[D]$ on merkitty mustalla paksulla viivalla.

4.4. Tulosten analysointi

4.4.1. Perusarvoinen tarkastelu

Kuvan 11 tilanne vastaa normaalia diagnostista päätöksentekotilannetta, jossa parametreihin ei ole liitetty epävarmuutta. Tällöin myöskään odotusarvoisiin hyötyihin ei liity epävarmuutta. Kuvasta 11 nähdään, että perusarvoilla tarkasteltuna SPECT-testi on aina PET-testiä parempi vaihtoehto. SPECT-testi siis dominoi PET-testiä koko välillä $P[D] \in [0, 1]$ minkä seurauksena PET-testiä ei tulisi ikinä valita sepelvaltimotautia tutkittaessa. Koska klinikko haluaa maksimoida odotusarvoisen hyödyn valitessaan testausstrategiaa, tulisi hänen valita potilaalle tehtävä toimenpide estimoimansa prioritodennäköisyyden perusteella, alla olevan kuvan 15 mukaisesti.

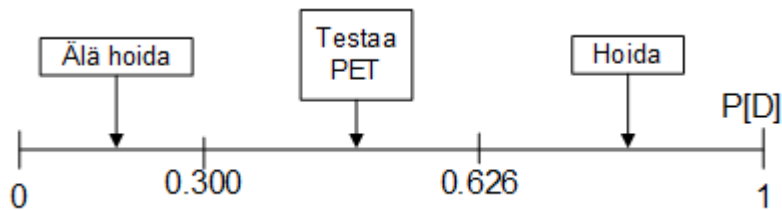


Kuva 15: Odotusarvoiset hyödyt maksimoivat testausstrategiat eri prioritodennäköisyyden väleillä, kun tarkastellaan odotusarvoisia hyötyjä parametrien perusarvoilla laskettuna.

Potilaalle ei siis tulisi tehdä toimenpiteitä, jos klinikko estimoii $P[D] \in [0, 0.367)$. Jos puolestaan estimoitu prioritodennäköisyys osuu välille $P[D] \in [0.367, 0.543)$, tulisi potilaalle tehdä SPECT – testi. Prioritodennäköisyyden osuessa välille $P[D] \in [0.543, 1]$, on odotusarvoiset hyödyt maksimoiva vaihtoehto aloittaa potilaan hoito. Vastaavat päätelmät voidaan tehdä myös muista tässä luvussa esitettävistä janakuvioista (kuvat 16-18).

4.4.2. Optimistinen tarkastelu

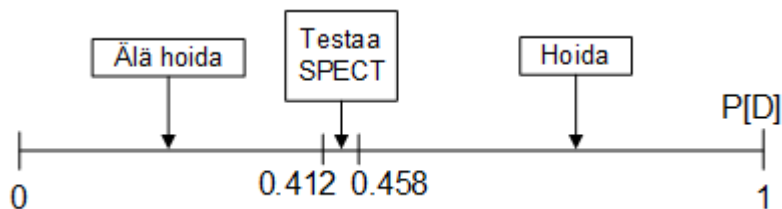
Parasta mahdollista tilannetta vastaa puolestaan kuva 12, jossa kaikki vaihtoehtoiset testausstrategiat saavat suurimmat mahdolliset odotusarvoisten hyötyjen arvot. Tätä tilannetta kuvaa ns. *maximax* –päättösääntö, jonka mukaan optimisti tekisi päätöksensä tämänkaltaisen tarkastelun perusteella [8]. Nyt PET-testin odotusarvoinen hyöty koko välillä $P[D] \in [0, 1]$ on suurempi kuin SPECT-testillä, minkä seurauksena PET dominoi SPECT-testiä. Maksimiarvoilla tarkasteltuna SPECT-testiä ei tulisi siis ikinä valita PET-testin sijasta. Kliinikon tulisikin valita testausstrategia estimoimansa prioritodennäköisyyden perusteella kuvan 16 mukaisesti. Huomattavaa on, että testausväli on laajentunut kuvan 15 esitykseen verrattuna.



Kuva 16: Odotusarvoiset hyödyt maksimoivat testausstrategiat eri prioritodennäköisyyden väleillä, kun tarkastellaan vain odotusarvoisten hyötyjen saamia maksimiarvoja.

4.4.3. Pessimistinen tarkastelu

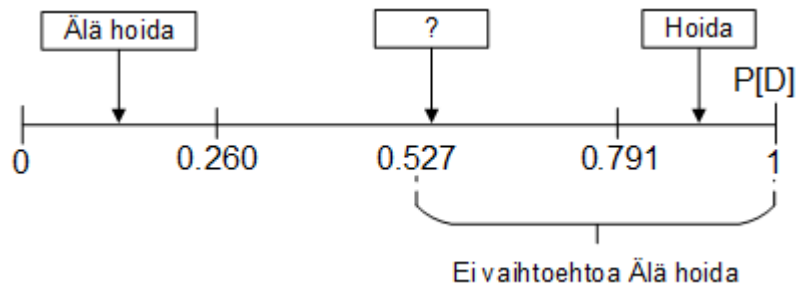
Kuva 13 vastaa pahinta mahdollista tilannetta, jossa kaikilla testausstrategioilla on huonoimmat mahdolliset odotusarvoisten hyötyjen arvot samanaikaisesti. Tätä tilannetta kuvaa ns. *maximin* – päätössääntö, jonka mukaan pessimisti tekisi päätöksensä tämänkaltaisen tarkastelun perusteella [8]. Huonoimmassa mahdollisessa tilanteessa SPECT-testi dominoi PET-testiä koko välillä $P[D] \in [0, 1]$, eikä PET-testiä tulisi ikinä valita näin tarkasteltuna. Pessimistisestä näkökulmasta tarkasteltuna klinikon tulisi prioritodennäköisyydestä riippuen valita testausstrategia alla olevan kuvan 17 mukaisesti. Kuvasta huomataan, että testausväli on kaventunut lähes olemattomaksi.



Kuva 17: Odotusarvot maksimoivat testausstrategiat eri prioritodennäköisyyden väleillä, kun tarkastellaan vain odotusarvoisten hyötyjen saamia minimiarvoja.

4.4.4. Kokonaisvaltainen tarkastelu

Kuva 14 esittää tilannetta, jossa kaikkien testausstrategioiden odotusarvoisten hyötyjen annetaan vaihdella parametrien epävarmuuksista (luku 3.3.) johtuvilla vaihteluvälillään. Tällöin vaihtoehto "Älä hoida" dominoi muita vaihtoehtoja välillä $P[D] \in [0, 0.260)$, ja vaihtoehto "Hoida" dominoi muita välillä $P[D] \in [0.791, 1]$. Siten näillä väleillä on aina järkevintä niitä vastaavat testausstrategiat. Näiden kahden strategian väliin jäävällä alueella $P[D] = P_0 \in [0.260, 0.791)$ ei kuitenkaan ole yhtä vaihtoehtoa, joka dominoisi kaikkia muita. Kuvasta 14 kuitenkin nähdään, että vaihtoehtoa "Älä hoida" dominoivat kaikki muut vaihtoehdot välillä $P[D] \in [0.527, 1]$, jolloin sitä ei kyseisellä välillä kannata valita. Kokonaisvaltainen tilanne on esitetty kuvassa 18.



Kuva 18: Odotusarvot maksimoivat testausstrategiat eri prioritodennäköisyyden väleillä, kun odotusarvoisten hyötyjen annetaan vaihdella. Kysymysmerkki (?) edustaa epävarmuutta testausstrategian valinnassa, sillä välillä $P[D] = P_0 \in [0.260, 0.791)$ ei ole yhtä dominoivaa testausstrategiaa, joka olisi aina muita parempi. Vaihtoehtoa "Älä hoida" ei kuitenkaan kannata valita välillä P_0 .

Koska välillä P_0 ei ole epävarmuuksien puitteissa yhtä parasta testausstrategiaa, voidaan asiaa tällä välillä tarkastella esimerkiksi jostain aiemmin esitetystä näkökulmasta: perusarvoisesta, optimistisesta tai pessimistisestä. Jos oletetaan, että perusarvoilla laskettu odotusarvoinen hyöty kuvaa myös odotusarvoisen hyödyn odotusarvoa, olisi testausstrategian valinnassa välillä P_0 keskimäärin järkevintä käyttää kuvan 15 esitystä. Vastaavasti positiivisesta näkökulmasta tarkasteluna tulisi strategioiden määrittelyssä tälle välille käyttää hyödyksi kuvaa 16, ja pessimistisestä näkökulmasta tarkasteltuna puolestaan kuvan 16 esitystä.

Käytännössä välin P_0 tarkastelemisessa tulisi ottaa myös huomioon mahdolliset käytännön seuraukset. Esimerkiksi kuvasta 14 nähdään, että testin SPECT vaihteluväli on kokonaisuudessaan PET-testin vaihteluvälin sisällä, mikä johtuu siitä, että PET-testin tarkkuuteen ja erottelukykyyneen liittyvä epävarmuus on paljon suurempaa kuin SPECT-testillä. Jos esimerkiksi jonkin tietyn kynnyksarvon, kuten $EU = 0.55$ alittaminen olisi kriittistä potilaan terveyden kannalta, tulisi klinikon valita pienemmän vaihteluvälin omaava SPECT-testi.

Koska parametreille ei ole määritelty erityisiä jakaumia, mutta niille on annettu tietyt välit perusarvon suhteen, voidaan kunkin parametrin perusarvon olettaa olevan myös kyseisen parametrin odotusarvo. Esimerkin tapauksessa tämä tarkoittaisi sitä, että perusarvoilla lasketut odotusarvoiset hyödyt ovat myös odotusarvoisten hyötyjen odotusarvoja $E[EU[\cdot]]$. Tällöin neutraalein vaihtoehto olisi käyttää perusarvoista näkökulmaa alueen P_0 tarkasteluun, minkä seurauksena päädyttäisiinkin lopulta kuvan 15 testausstrategioihin prioritodennäköisyyden funktiona.

5. Yhteenveto ja pohdinta

Työssä tutkittiin, miten diagnostisiin testeihin ja potilaan sairauden tilaan liittyvä epävarmuus vaikuttaa diagnostisen testausstrategian valinnassa. Ensin tarkasteltiin, miten kliinikot estimoivat potilaan sairauden prioritodennäköisyyksiä käytännössä. Tämän jälkeen tutkittiin, miten diagnostisten testien antaman lisäinformaation avulla voidaan laskea testin jälkeinen posterioritodennäköisyys. Sen jälkeen tarkasteltiin päätöspuita sekä odotusarvoista hyötyä, ja perusteltiin epävarmuuksien mallintaminen eri parametreille. Lopuksi tutkittiin esimerkin avulla epävarmuuden vaikutuksia testausstrategian valintaan.

Epävarmuuksien huomioiminen vaikutti erityisesti siihen, ettei kaikkia muita vaihtoehtoja dominoivaa testausstrategiaa aina löytynyt tietyille prioritodennäköisyyden väleille. Tällöin valittavaa strategiaa voitiin kuitenkin tutkia dominanssien avulla perusarvoisesta, optimistisesta tai pessimistisestä näkökulmasta. Ensimmäinen näistä vastasi niin sanottua perustilannetta, kaksi jälkimmäistä puolestaan parasta ja huonointa mahdollista tilannetta. Täysin toisistaan poikkeavista kuvista 15 – 17 nähtiin, että epävarmuudella oli merkittävä vaikutus optimaalisen testausstrategian valinnassa. Nämä vaikutukset näkyivät niin valittavissa diagnostisissa testeissä (PET tai SPECT riippuen tarkastelunäkökulmasta) kuin prioritodennäköisyyden väleissäkin.

Esimerkin yhteydessä huomattiin testien tarkkuuden ja erottelukyvyn vaikuttavan siihen, kuinka pienillä tai suurilla prioritodennäköisyyden arvoilla potilas kannatti lähettää diagnostisiin testeihin. Jos diagnostisen testin tarkkuus ja erottelukyky olivat korkeat, testi antoi niin paljon lisäinformaatiota sairauden tilasta, että potilas kannatti lähettää testeihin laajemmalla prioritodennäköisyyden välillä kuin tarkkuuden ja erottelukyvyn ollessa matalat. Esimerkiksi PET-testin tarkkuus ja erottelukyky olivat pessimistisestä näkökulmasta tarkasteltuna niin matalat, ettei sitä olisi ollut järkevää suorittaa missään tapauksessa.

Liitteessä B esitetyn laskennan perusteella huomattiin hyötyjen estimoinnin haasteellisuus. Todellisuudessa hyötyjen estimoinnissa tulisi ottaa liitteessä läpikäytyjen näkökulmien lisäksi huomioon myös monia muita seikkoja, minkä vuoksi tarkan estimaatin saaminen onkin käytännössä erittäin haasteellista. Hyötyestimaattien epävarmuuden mallinnus onkin erittäin perusteltua.

Jatkotarkastelujen kannalta epävarmuuksien tarkastelussa voisi ottaa huomioon parametrien jakaumat, ja niiden avulla saatavan odotusarvoisen hyödyn jakauma. Strategian valintaa voisi tarkastella myös tutkimalla osittain päällekkäin menevien strategioiden rajaamia pinta-aloja. Lisäksi valinnassa voisi huomioida populaation jakauman prioritodennäköisyyden suhteen, jolloin testausstrategioiden odotusarvoisia hyötyjä voisi tarkastella koko populaation näkökulmasta.

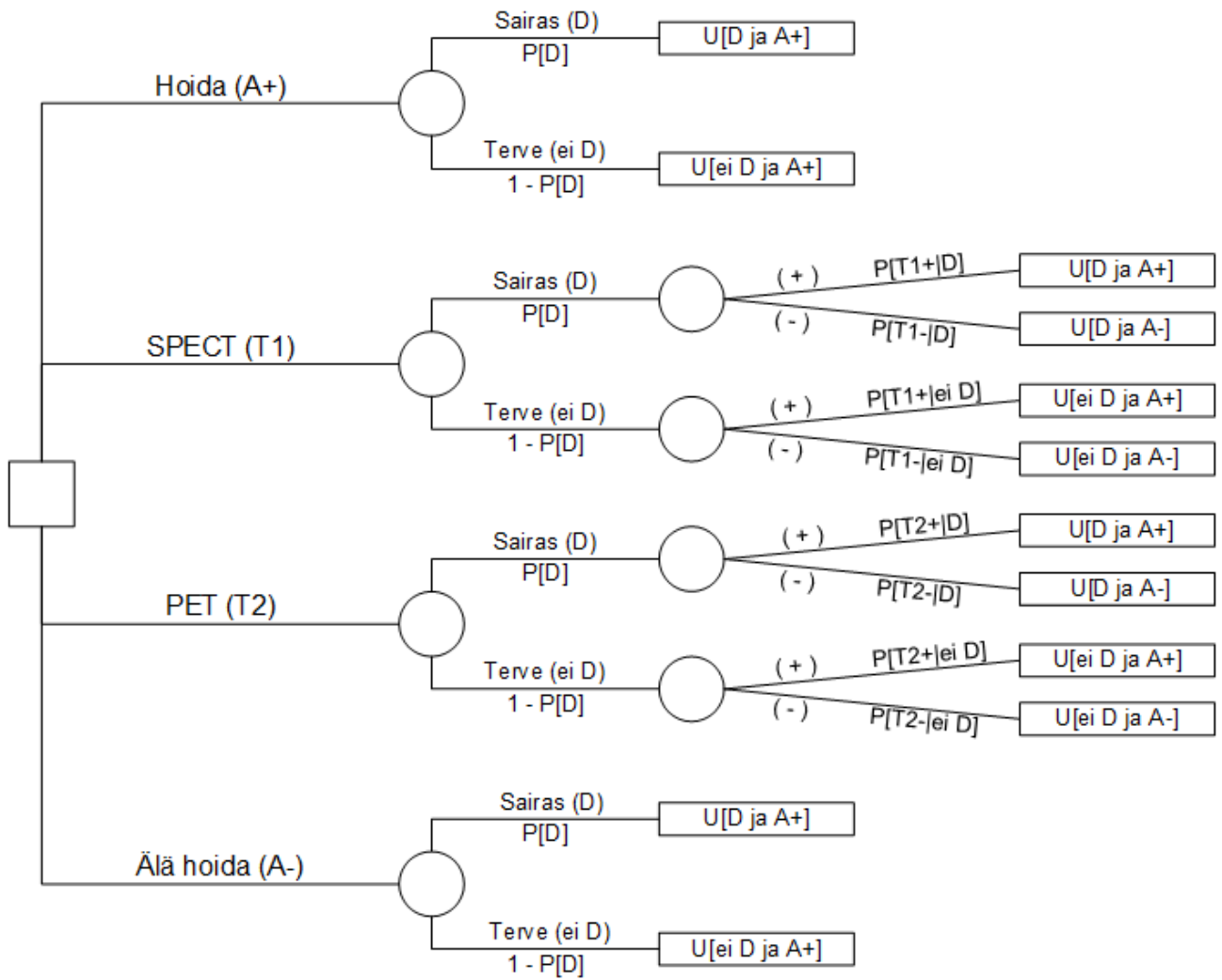
Kaiken kaikkiaan epävarmuuden huomioiminen olisi tärkeää sekä potilaan että yhteiskunnan kannalta, sillä kliinikon tekemällä strategiavalinnalla on vaikutuksia näihin molempiin. Potilaan kannalta olisi tärkeää tietää esimerkiksi testauksesta koituvat pahimmat mahdolliset seuraukset. Yhteiskunnan kannalta olisi puolestaan järkevintä hoitaa sekä sairaus että sen testaus mahdollisimman kustannustehokkaasti. Terveystieteiden huollossa ollaankin siirtymässä entistä kustannustehokkaampaan ja arvo-painotteisempaan suuntaan, jossa operaatiotutkimuksella tulee olemaan suuri merkitys [15].

Viitteet

- [1] Sox HC, Higgings MC, Owens DK. *Medical Decision Making*, 2nd edition. 2013. John Wiley & Sons Ltd.
- [2] Raiffa H. *Decision Analysis: Introductory Lectures on Choices Under Uncertainty*. 1968. Addison-Wesley.
- [3] Simel DL, Rennie D. *The Rational Clinical Examination: Evidence-based Clinical Decision-Making*, 4th edition. 2008. McGraw-Hill Medical.
- [4] McNeil BJ, Adelstein SJ. *Determining the Value of Diagnostic and Screening Tests*. Journal of Nuclear Medicine. 1976;17:439-448.
- [5] Felder S, Mayrhofer T. *Medical Decision Making: A Health Economic Primer*. 2011. Springer.
- [6] Bossuyt PMM, Reitsma JB, Linnet K, Moons KGM. Beyond Diagnostic Accuracy: The Clinical Utility of Diagnostic Tests. Clinical Chemistry. 2012;58:1636-1649.
- [7] Deeks J. *Systematic Reviews of Evaluations of Diagnostic and Screening Tests*. BMJ. 2001;323:157-62.
- [8] Liesiö J, Mild P, Salo A. *Preference Programming for Robust Portfolio Modeling and Project Selection*. European Journal of Operational Research. 2007;181:1488-1505.
- [9] Kalyanmoy D. *Multi-Objective Optimization Using Evolutionary Algorithms*. 2002. Wiley.
- [10] Parker MW, Iskandar A, Limone B, Perugini A, Kim H, Jones C, Calamari B, Coleman CI, Heller GV. *Diagnostic Accuracy of Cardiac Positron Emission Tomography Versus Single Photon Emission Computed Tomography for Coronary Artery Disease: A Bivariate Meta-Analysis*. Circulation by American Heart Association: Cardiovascular Imaging. 2012;5:700-707. Saatavilla: <http://circimaging.ahajournals.org/content/5/6/700>. Luettu 23.10.2014.
- [11] Villines TC, Min JK. *Comparing Outcomes and Costs Following Cardiovascular Imaging*. Journal of the American College of Cardiology. 2014;63(10):1009-1010.
- [12] Hall H. *Article: Dr. Amen's Love Affair with SPECT Scans*. Science-Based Medicine. 2013. Saatavilla: <http://www.sciencebasedmedicine.org/dr-amens-love-affair-with-spect-scans/>. Luettu 13.11.2014.
- [13] Suomen Säteilyturvakeskus (STUK). *Säteilyn terveystvaikutukset*. Elokuu 2009. Saatavilla: http://www.stuk.fi/sateily-ymparistossa/fukushima/fi_FI/fukushima_12-08-2011/_files/89705092504300839/default/katsaus_sateilyn_terveysvaikutukset_elokuu_2009.pdf. Luettu 13.11.2014.
- [14] Terveyskirjasto, Lääkärikirja Duodecim, kardiologian erikoislääkäri Raimo Kettunen. *Sepelvaltimotauti*. Saatavilla: http://www.terveyskirjasto.fi/terveyskirjasto/tk.koti?p_artikkeli=dlk00077. Luettu 13.11.2014.

- [15] Gray M. *Value: Operations Research and the New Health Care Paradigm*. Operations Research for Health Care Journal. 2012:1;20-21.

Liite A – Esimerkin testausstrategioiden päätöspuu



Liite B – Esimerkinomainen hyötyjen laskeminen

Testien nettohyötyjä $U[SPECT]$ ja $U[PET]$ voidaan arvioida sepelvaltimotaudin testaukseen liittyvän SPARC-tutkimuksen [11] avulla, jossa testejä on vertailtu. Ajatellaan testin T nettohyötyjä välillä $U[T] \in [-1, 1]$, missä $-1 = \text{"äärimmäinen haitta"}$, $0 = \text{"ei haittaa/hyötyä"}$ ja $1 = \text{"äärimmäinen hyöty"}$. SPARC-tutkimuksen mukaan PET-testin suorittaminen on 22 % kalliimpaa kuin SPECT:n eli taloudellisesta näkökulmasta pätee $U[PET]_t = 1.22 \times U[SPECT]_t$. Arvioidaan kolmen vuoden palkkakustannuksien $(3000 \frac{\text{€}}{\text{kk}} \times 3 \times 12 \text{ kk} = 108\,000 \text{ €})$ menetystä äärimmäisenä haittana potilaalle. Koska yhden SPECT-testin suorittaminen maksaa $\$3500 \approx 2800 \text{ €}$ [12], saadaan SPECT-testin nettohyödyksi talouden suhteen arvioitua $U[SPECT]_t = \frac{2800 \text{ €}}{108\,000 \text{ €}} * (-1) \approx -0.0259$. Tällöin $U[PET]_t = -0.0316$.

Artikkelin [12] mukaan myös potilaalle aiheutunut kokonaissäteilyannos (suurempi määrä haitallisempi) PET-testistä on 2.0 mSV ($SV = \text{Sievert}$), kun taas SPECT-testin vastaava luku on 0.6 mSV . Näin ollen PET-testin suorittaminen olisi lineaarisesti tarkasteltuna säteilyn kannalta noin kolme kertaa haitallisempi potilaalle kuin SPECT, jolloin $U[PET]_r = 3 \times U[SPECT]_r$. Koska Säteilyturvakeskuksen mukaan [13] kuolettava annos on $10\,000 \text{ mSV}$ (äärimmäinen haitta potilaalle), voidaan ajatella PET-testin säteilystä aiheutuvan nettohyödyn olevan $U[PET]_r = \frac{2.0 \text{ mSV}}{10\,000 \text{ mSV}} \times (-1) = -2 \times 10^{-4}$, ja SPECT-testille puolestaan $U[SPECT]_r \approx -6.67 \times 10^{-5}$.

Talous- ja säteilynäkökulmat yhdistämällä nettohyöty testille PET on $U[PET] = U[PET]_r + U[PET]_t \approx -0.0318$, ja vastaavasti $U[SPECT] = U[SPECT]_r + U[SPECT]_t \approx -0.0260$. Edelliset approksimaatiot ottavat huomioon vain potilaan suhteen testiin. Approksimoidaan yhteiskunnallisten vaikutusten olevan yhtä suuret, eli kerrotaan hyödyt kahdella, jolloin $U[SPECT] = -0.052$ ja $U[PET] = -0.064$. Nämä approksimaatiot ovat karkeita, joten annetaan niiden vaihdella 5 % suuntaansa. Lopputulokset väleinen on esitetty taulukossa 6.

Viimeisenä arvioidaan esimerkinomaisesti toimenpide-terveydentila-kombinaatioiden (hoida/älä hoida, sairas/ei sairas) hyödyt väliltä $[0, 1]$ käyttämällä apuna Terveyskirjaston kuvausta sepelvaltimotaudista [14]. Koska $U[D \text{ ja } A+]$ merkitsee sepelvaltimotaudin omaavan henkilön hoitamista, on hoidosta saatava hyöty suuri. Hoito voi kuitenkin erittäin pienellä todennäköisyydellä johtaa kuolemaan, joten arvioidaan $U[D \text{ ja } A+] \approx 0.95$. Jos sepelvaltimotaudin omaavaa potilasta ei hoideta, potilas voi tästä huolimatta parantua sairaudesta. Sairaus kuitenkin voi uusiutua helposti, joten approksimoidaan, että noin 10 % potilaista voi parantua kokonaan ja loput kuolee aikaisemmin, jolloin $U[D \text{ ja } A-] \approx 0.10 \times 1 + 0.90 \times 0 \approx 0.10$. Jos tautia ei ole, mutta potilasta hoidetaan, annettu lääkitys vaikuttaa potilaaseen. Lääkkeet eivät kuitenkaan ole hengenvaarallisia, joten arvioidaan, että 50 % potilaista kokee vaikutuksen 0.50 (haitallinen), 40 % kokee vaikutuksen 0.10 (erittäin haitallinen) ja 10 % kokee vaikutuksen 0.75 (hieman haitallinen) eli $U[\text{ei } D \text{ ja } A+] \approx 0.5 \times 0.5 + 0.4 \times 0.10 + 0.1 \times 0.75 \approx 0.37$. Jos tautia ei ole, eikä potilasta hoideta, ei tästä koidu haittaa potilaalle, joten $U[\text{ei } D \text{ ja } A-] \approx 1$. Annetaan myös näiden approksimaatioiden vaihdella 5 % suuntaansa. Sairauden hyödyt väleinen on esitetty taulukossa 7.