

# 6

## JOHDATUS EPIDEMIOLOGIAAN

Anssi Auvinen

### SISÄLLYSLUETTELO

6.1	Yleistä .....	78
6.2	Epidemiologiset tutkimusasetelmat .....	82
6.3	Riskimallit .....	86
6.4	Annosvaste .....	89

## 6.1 | Yleistä

Epidemiologia on tutkimusala, jonka kohteena on sairauksien esiintymisen väestössä. Perinteisen määritelmän mukaan epidemiologia tutkii sairauksien esiintymistä ja siihen vaikuttavia tekijöitä. Viime aikoina on kuitenkin korostettu, että epidemiologia kattaa myös muun muassa sairauksien ennusteen, kliinisen toiminnan ja terveydenhuollon tutkimuksen. Tämän perusteella epidemiologia on määritelty lääketieteelliseksi esiintyvyystudkimukseksi tai lääketieteelliseksi syytutkimukseksi. Epidemiologian historiallinen tausta on 1800-luvun epidemioiden tutkimuksessa, mutta nykyisin pääosa tutkimuksesta liittyy ei-tarttuviin sairauksiin kuten syöpään sekä sydän- ja verenkiertoelimistön sairauksiin.

Epidemiologiseksi transitioksi kutsutaan yhteiskunnan kehityksen myötä länsimaissa tapahtunutta sairauskirjon muutosta. Muutoksen myötä keskeisin terveysongelma on vaihtunut lasten infektiokuolleisuudesta vanhojen ikäluokkien ”elintamosairauksiksi”. Epidemiologian keskeisiä saavutuksia 1900-luvulla olivat muun muassa tupakan ja syövään välisen yhteyden löytäminen, menestyksekkäiden rokotusohjelmien kehittäminen, sekä elintapojen ja sydän- ja verisuonitautien yhteyden selvittäminen.

Epidemiologialle ominaisia piirteitä ovat empiirisuus, kvantitatiivisuus sekä probabilistinen lähestymistapa. Empiirisuus tarkoittaa sitä, että epidemiologia perustuu suunnitelmalliseen havaintojen keruuseen. Kvantitatiivisuus ja probabilistisuus tarkoittavat sitä, että havainnot käsitellään numeerisin menetelmin. Käsitelystä materiaalista voidaan tehdä todennäköisyyttä koskevia päätelmiä. Myös epidemiologinen syykäsitely perustuu todennäköisyyteen: syytekijä muuttaa sairauden esiintymisen todennäköisyyttä.

Epidemiologian tärkeimpiä lähitieteitä ovat lääketiede, tilastotiede ja demografia. Lääketieteen yhteys epidemiologiaan on ilmeinen, sillä molemmat käsittelevät terveyttä ja sairautta. Epidemiologia käsittelee näitä ilmiöitä tyyppillisesti väestötasolla (kansanterveys), kun taas kliinisessä lääketieteessä painopiste on yksilöllä. Näyttöön perustuva lääketiede on esimerkki ajattelutavasta, joka yhdistää molempia lähestymistapoja. Tilastotiede on epidemiologin työkalu, sillä epidemiologisten aineistojen käsittelyssä käytettävät menetelmät perustuvat tilastotieteeseen. Demografia eli väestötiede tutkii väestöjen dynamiikkaa kuten syntyvyyttä ja kuolleisuutta. Monet epidemiologiset tutkimusasetelmat ja tunnusluvut ovat peräisin demografiasta.

Tutkimusväestön määrittely on luonnollisesti hyvin tärkeää, koska epidemiologia tutkii terveyteen liittyviä tapahtumia väestössä. Tutkimusväestö määri-

tellään sekä ajan että paikan suhteen, ja siihen kuuluvat ihmiset pyritään identifioimaan. Suurten tutkimusten toteuttamista auttavat väestörekisterit, jotka ovat parhaassa tapauksessa kattavuudeltaan lähes täydellisiä ja mahdollistavat yksilöiden tarkan identifioinnin henkilötunnuksen avulla. Tarkkaan määrittelystä väestöstä voidaan valita tarkoituksenmukainen otos tutkimusta varten. Tyypillisesti valitaan satunnaisotos, mutta joissain tapauksissa myös kokonaisotos eli kaikki tietyt ehdot täyttävät henkilöt. Suomi ja muut Pohjoismaat edustavat useissa suhteissa epidemiologisen tutkimuksen kärkeä. Yhtenä tärkeänä seikkana tähän on vaikuttanut korkeatasoisten rekistereiden toiminta. Väestörekistereiden lisäksi Suomessakin toimii useita sairausrekistereitä, joista mainittakoon vuonna 1952 perustettu Syöpärekisteri ja STAKESin ylläpitämä sairaaloiden hoitoilmoitusrekisteri.

Sairauden esiintyvyyttä kuvataan epidemiologiassa yleisimmin ilmaantuvuuden avulla. Ilmaantuvuus eli insidenssi määritellään uusien tapausten määräksi suhteessa tutkimusväestön kokoon. Yleensä ilmaantuvuusluku ilmaistaan suhteutettuna väestöaikaan (kuten henkilövuosiin) eli henkilöiden määrän ja seuranta-ajan funktiona. Harvinaisten sairauksien, kuten syövän ilmaantuvuus ilmaistaan yleensä 100 000 henkilövuotta kohti. Vakioitu ilmaantuvuusuhde (standardised incidence ratio, SIR) kuvaa ilmaantuvuutta suhteessa iän, sukupuolen ja tarkastelujakson osalta vastaavaan väestöön. SIR ilmaistaan havaittujen tapausten suhteena odotettuun tapausmäärään. Arvo 1 osoittaa ilmaantuvuuden vastaavan vertailuväestöä, sitä suurempi luku suurentunutta vaaraa, ja pienempi suojavaikutusta. Kuolleisuus vastaavasti on kuoleman ilmaantuvuutta, ja vakioitu kuolleisuussuhde (standardised mortality ratio, SMR) kuvaa kuolleisuutta suhteessa muuhun väestöön.

Toinen tapa kuvata sairauden esiintymistä on vallitsevuus eli prevalenssi, joka ilmaisee sairaustilan yleisyyttä, kun taas ilmaantuvuus kuvaa sairastumisten yleisyyttä. Vallitsevuus on sairaiden osuus tutkimusväestöstä ja se voidaan ilmoittaa joko tietyllä hetkellä (pisteprevalenssi) tai tietyllä jaksolla (periodiprevalenssi). Vallitsevuutta voidaan käyttää sairauden yleisyyden kuvaamiseen erityisesti silloin kun tarkkaa sairastumishetkeä on vaikea tai mahdoton määrittää. Esimerkkejä tällaisista vaikeasti ajallisesti määriteltävistä ovat sikiöepämuodostumat, HIV-infektio ja kohonnut verenpaine. Sairauden syytekijöiden tutkimiseen vallitsevuus soveltuu huonosti, sillä se riippuu paitsi ilmaantuvuudesta, myös kuolleisuudesta ja parantuvuudesta.

Niitä tekijöitä, joiden mahdollista vaikutusta selvitetään, kutsutaan epidemiologiassa yleisnimellä altistukset (exposure). Termin käyttö poikkeaa yleiskielestä, sillä epidemiologiassa altistus voi olla fyysikaalisen tai kemi-

allisen tekijän ohella esimerkiksi demografinen, sosiaalinen tai biologinen ominaisuus. Yleisimmin tutkittuihin altisteisiin lukeutuvat muun muassa ikä, sukupuoli, elintapa, ammatti, sosiaaliluokka, asuinalue, sairaus, lääkitys ja geneettinen ominaisuus. Tutkittavaa sairautta tai terveydentilaa kutsutaan vasteeksi (outcome). Epidemiologisessa tutkimuksessa kerätään systemaattisesti tietoa sekä altistuksesta että vasteesta. Toisin sanoen tehdään altistusta ja vastetta koskevia mittauksia, joskin 'mittaus' saattaa perustua esimerkiksi kyselylomakkeeseen tai kuolintodistukseen.

Mittauksen kykyä tunnistaa mitattava ominaisuus sanotaan validiteetiksi. Validiteetti voidaan jakaa herkkyydeksi (sensitiivisyys) eli kyvyksi havaita ominaisuuden olemassaolo ja tarkkuudeksi (spesifisyys) eli kyvyksi havaita ominaisuuden puuttuminen. Mittauksiin liittyy vääjäämättä jonkinasteisia virheitä. Virheen lähteenä voi olla menetelmä, mittaustapa, tulosten kirjaus tai biologinen vaihtelu. Jos vertailtavissa ryhmissä on samansuuruisia satunnaisvirheitä, tämä laimentaa havaittuja yhteyksiä. Erisuuruisen satunnaisvirheen puolestaan aiheuttaa harhaa ja saattaa vääristää tuloksia joko korostamalla tai peittämällä todellisia yhteyksiä. Jos mitattava väestö poikkeaa kohdeväestöstä mitattavan tekijän suhteen, puhutaan valikoitumisharhasta. Jos taas samasta ominaisuudesta saadaan tietoa vertailtavista ryhmistä eri lailla, puhutaan informaatioharhasta.

Ennen kuin voidaan tehdä päätelmiä tutkittavan tekijän vaikutuksesta sairauden esiintyvyyteen, epidemiologisessa tutkimuksessa on arvioitava satutuman, harhan ja sekoittuneisuuden mahdollista vaikutusta. Satunnaisvirheen osuutta voidaan arvioida laskemalla tilastollinen merkitsevyys tai luottamusväli. Tilastollinen merkitsevyys kertoo kuinka suurella todennäköisyydellä sattuma tuottaisi saman suuruisen eron, mikäli todellisuudessa ryhmien välillä ei ole eroja. Luottamusväli ilmaisee sen alueen, millä todellinen arvo on halutulla todennäköisyydellä, esimerkiksi 95 prosenttia. Satunnaisvirhe pienenee otoskoon kasvaessa, kun taas systemaattinen virhe, harha, on riippumaton otoskoosta. Tilastollisen merkitsevyyden tulokinnassa on kuitenkin otettava huomioon myös biologinen merkitsevyys: suuressa aineistossa häviävän pienetkin, käytännössä merkityksettömät erot voivat olla tilastollisesti merkitseviä. Toisaalta pienessä tutkimuksessa edes huomattava vaikutus, esimerkiksi moninkertainen syöpävaara ei välttämättä ole tilastollisesti merkitsevä. Onkin suositeltavaa jo ennen tutkimuksen aloittamista tehdä voimalaskelmat, joiden perusteella voidaan arvioida, onko otoskoko sopiva realistisesti odotettavissa olevan vaikutuksen tutkimiseen.

Harha aiheutuu siitä, että vertailtavat väestöt eivät ole vertailukelpoisia. Valikoitumisharha syntyy, jos osallistujat poikkeavat tutkittavien

tekijöiden osalta kohdeväestöstä (niistä, jotka oli valittu tutkimukseen). Kohorttitutkimuksessa tämä tarkoittaa sitä, että sairauden todennäköisyys liittyy osallistumisen todennäköisyyteen tai tapaus-verrokkitutkimuksessa altistuksen todennäköisyys liittyy osallistumisen todennäköisyyteen. Tutkimushenkilöiden ei kuitenkaan tarvitse edustaa koko väestöä, vaan tutkijan määrittelemää kohdeväestöä. Informaatioharha puolestaan syntyy, jos tutkittavista saatu informaatio ei ole vertailukelpoista, esimerkiksi tiedon tarkkuus poikkeaa yli- tai aliraportoinnin vuoksi. Asbestille altistuneiden keuhkomuutoksia saatetaan seurata tarkemmin kuin altistumattomien, mikä suurentaa sairauden havaitsemisen todennäköisyyttä. Vastaavasti tapaus-verrokkitutkimuksessa vakavasti sairastuneet potilaat saattavat antaa verrokkeja tarkempia tietoja altistushistoriastaan, koska ovat motivoituneita tutkimukseen ja haluavat löytää syyn sairastumiselleen.

Selvitettävän tekijän lisäksi epidemiologisessa tutkimuksessa on usein muita, ulkoisia tekijöitä, jotka vaikuttavat sairauden esiintymiseen. Näiden vaikutusta pyritään kontrolloimaan, etteivät ne aiheuta sekoittuneisuutta. Sekoittuneisuus (confounding) johtuu siitä, että muut tekijät korreloivat tutkittavan altistuksen kanssa. Tällöin vertailtavat ryhmät poikkeavat toisistaan muiden sairauden esiintymiseen vaikuttavien tekijöiden suhteen. Vaikutusta voidaan vakioida sitä tehokkaammin mitä tarkemmat ja yksityiskohtaisemmat tiedot on käytettävissä. Niiden tekijöiden vaikutusta ei voida vakioida, joista ei ole kerätty lainkaan tietoa. Sekoittuneisuutta voidaan kontrolloida vakioimalla tai rajaamalla aineisto ulkoisen tekijän suhteen, laskemalla vakioituja esiintymislukuja tai käyttämällä tilastollisia monimuuttujamenetelmiä.

Muovautuneisuus (effect modification) tarkoittaa vaikutuksen vaihtelua muovaavan tekijän suhteen. Muovautuneisuus voi olla antagonismia, jossa kahden tekijän (a, b) yhteinen vaikutus on pienempi kuin niiden riippumaton vaikutus ( $\lambda_{a,b} < \lambda_a, \lambda_b$ ) tai synergismia, jolloin tekijöiden yhteisvaikutus on suurempi kuin niiden erilliset vaikutukset ( $\lambda_{a,b} > \lambda_a, \lambda_b$ ). Muovautuneisuus ei kuitenkaan edellytä, että muovaavalla tekijällä olisi itsenäinen vaikutus. Yksi esimerkki muovautuneisuudesta on säteilyn aiheuttaman syöpävaaran suuruuden vaihtelu iän mukaan. Vuorovaikutus eli interaktio on tekninen nimitys muovautuneisuudelle. On huomattava, että muovautuneisuus riippuu käytetystä riskimallista: jos kahden tekijän vaikutus additiivisessa mallissa (katso sivu 88) on synergistinen, multiplikatiivisessa mallissa muovautuneisuutta ei välttämättä havaita. Ääritapauksissa se voi olla multiplikatiivisessa mallissa antagonistinen tai synergistinen.

## 6.2 | Epidemiologiset tutkimusasetelmat

Tavoitteidensa perusteella epidemiologiset tutkimukset voidaan jakaa kuvaileviin ja analyttisiin. Kuvailevan tutkimuksen tulokset ovat esiintyvyyksilukuja, kuten ilmaantuvuus, kuolleisuus ja/tai vallitsevuus. Kuvailevan tutkimuksen asetelma on yleensä poikkileikkaustutkimus eli yhdellä hetkellä määritelty satunnaisotos siitä väestöstä, jota koskevia päätelmiä aiotaan tehdä. Tyypillisesti kuvailevassa tutkimuksessa kerätään tietoja vain vastetapahtumasta. Kuitenkin raja kuvailevan ja analyttisen tutkimuksen välillä on epäselvä, sillä usein kuvailevassakin tutkimuksessa käytetään selittävinä tekijöinä esimerkiksi ikää ja sukupuolta.

Analyttisessä epidemiologiassa tutkimuksen kohteena on altistuksen ja vasteen yhteys. Tavoitteet voivat liittyä hypoteesin testaukseen, jolloin pyritään aineiston perusteella päättämään, pitääkö testattava hypoteesi paikkansa. Hypoteesin testausta varten muodostetaan kaksi vaihtoehtoista väitettä: nollahypoteesin ( $H_0$ ) mukaan tutkittavalla tekijällä ei ole vaikutusta, ja vaihtoehtoisen eli tutkimushypoteesin ( $H_1$ ) mukaan sillä on vaikutus. Hypoteesin testaus on laadullista päättelyä tilastollisen merkitsevyyden perusteella. Sen lähtökohdana oletetaan  $H_0$ :n pitävän paikkansa. Tilastollisen testauksen perusteella arvioidaan kuinka suurella todennäköisyydellä tutkimusaineistossa havaittu - tai suurempi - ero syntyisi, mikäli  $H_0$  on tosi.  $H_0$  hylätään ja tutkimushypoteesi hyväksytään, mikäli tämä todennäköisyys eli tilastollinen merkitsevyys on riittävän pieni. Raja-arvona on yleensä 0,05 eli 5 prosenttia. Vaihtoehtoisesti tavoite voi olla riskinarviointi eli estimointi, joka pyrkii määrälliseen arvioon. Tällöin arvioidaan vaikutuksen suuruutta (esiintyvyyssuhdetta tai -eroa) ja lasketaan sille luottamusväli, joka kuvaa kuinka suuri todellinen vaikutus on määritellyllä todennäköisyydellä - yleensä 0,95 eli 95 prosenttia. Luottamusväli lasketaan tietulle jakaumalle satunnaisvirheen perusteella.

Analyttisen epidemiologian tyypillisiä tutkimusasetelmia ovat ekologinen tutkimus, tapaus-verrokkitutkimus, kohorttitutkimus ja satunnaistettu koe.

### Ekologinen tutkimus

Ekologisessa tutkimuksessa tiedot altistuksesta ja vasteesta ovat käytettävissä vain ryhmätasolla ilman, että tutkija itse pystyisi ryhmittelemään havaintoja. Ekologiset tutkimukset vertailevat tyypillisesti altistuksen ja vasteen esiintymistä eri alueiden tai ajanjaksojen välillä. Ekologisen tutkimuksen heikkouksia ovat ekologinen harha, satunnaisvirheen eroja ko-

rostava vaikutus ja sekoittavien tekijöiden kontrollin vaikeus. Ekologinen harha johtuu siitä, että tiedot ovat saatavissa vain ryhmätasolla, eikä tiedettä missä altistusryhmässä sairaustapaukset esiintyvät ryhmän sisällä.

## Tapaus-verrokkitutkimus

Tapaus-verrokkitutkimuksessa määritellystä tutkimusväestöstä valitaan kaikki sairaustapaukset määritellyltä aikaväliltä. Tapaus-verrokkitutkimusta kutsutaan myös sairauslähtöiseksi tutkimukseksi. Vertailuryhmäksi valitaan samasta väestöstä otos niistä henkilöistä, joilla ei määritellyllä hetkellä ollut sairautta. Verrokkit voidaan valita joko satunnaisesti kaikkien itapausten joukosta tai kaltaistettuna siten, että verrokkien jakauma vastaa joidenkin ominaisuuksien suhteen tapauksia. Kaltaistuksella pyritään varmistamaan tapauksien ja verrokkien vertailukelpoisuus. Vertaamalla tapauksien ja verrokkien altistumista voidaan tehdä päätelmiä altistuksen ja sairauden yhteydestä. Altistuksen vaikutusta mitataan tapaus-verrokkitutkimuksessa ristitulosuhteella (odds ratio, OR). Katso taulukko 6.1.

	Altistunut	Altistumaton
Tapaus	A	B
Verrokki	C	D

**TAULUKKO 6.1** Kaltaistamattoman tapaus-verrokkitutkimuksen analyysi

Taulukossa A kuvaa altistuneiden tapauksien määrää, B altistumattomien tapauksien määrää, C altistuneiden verrokkien määrää ja D altistumattomien verrokkien määrää. Ristitulosuhde,  $OR = a/b : c/d = ad/bc$

Tapaus-verrokkitutkimuksella voidaan selvittää useiden tekijöiden vaikutusta samaan sairauteen, ja se on tehokkain tapa tutkia harvinaisia sairauksia. Tapaus-verrokkitutkimuksen tilastollinen voima riippuu tutkimusväestön koon lisäksi altistusjakaumasta. Yleensä tilastollinen voima on suurin, jos altistus on tasaisesti jakautunut joko tapauksien tai verrokkien joukossa. Pienen tapausmäärän rajoittamaa tilastollista voimaa voidaan jossain määrin kasvattaa lisäämällä kaltaistussuhdetta eli verrokkien määrää suhteessa tapauksiin. Lisäverrokkien tuoma hyöty kuitenkin pienenee nopeasti kaltaistussuhteen kasvaessa. Kaltaistetun tutkimuksen analyysi poikkeaa kaltaistamattomasta, sillä tapaus-verrokkipari – tai useammista tutkimushenkilöistä koostuva tapaus-verrokkiseti – analysoidaan yhtenä yksikkönä (katso taulukko 6.2).

Tapaus-verrokkitutkimuksessa on oleellista määritellä kattavasti tutkimusväestö. Lisäksi on pystyttävä identifioimaan sekä keräämään tietoa kaikista tapauksista. Verrokki täytyy pystyä valitsemaan altistuksesta riippumatta. Poikkeaminen näistä periaatteista aiheuttaa helposti valikoitumisharhaa. Useimmissa tapaus-verrokkitutkimuksissa altistustiedot kerätään tapauksilta vasta sairauden toteamisen jälkeen. Tässä tilanteessa diagnoosi voi vaikuttaa altistustietojen saamiseen ja aiheuttaa muistiharhaa. Tyypillisesti muistiharhaa syntyy siten, että tapaukset ovat paremmin motivoituneita ja raportoivat siksi altistushistoriansa täydellisemmin, erityisesti jos on kyse tavanomaisista asioista kuten esimerkiksi röntgenkuvista. Verrokkien aliraportointia ei tällöin voida arvioida pelkästään varmistamalla raportoidut tiedot muista lähteistä, vaan oleellista olisi löytää myös raportoimattomat tiedot. Altistustietojen kerääminen tutkimushenkilöistä riippumattomista lähteistä estää tämän harhan mahdollisuuden.

		Verrokki	
		Altistunut	Altistumaton
Tapaus	Altistunut	A	B
	Altistumaton	C	D

**TAULUKKO 6.2** Kaltaistetun (1:1) tapaus-verrokkitutkimuksen analyysi

Taulukossa A kuvaa sellaisten tapaus-verrokkiparien määrää, joissa sekä verrokki että tapaus ovat altistuneet. B kuvaa sellaisten parien määrää, joissa tapaus on altistunut ja verrokki altistumaton. C kuvaa sellaisten parien määrää, joissa tapaus on altistumaton ja verrokki altistunut. D kuvaa sellaisten tapausten määrää, joissa molemmat ovat altistumattomia.

Ristitulosuhde,  $OR=b/c$

## Kohorttitutkimus

Kohorttitutkimuksessa muodostetaan vertailtavat ryhmät altistuksen perusteella ja sitä kutsutaankin altistuslähtöiseksi tutkimukseksi. Kohorttitutkimuksessa määritellään ensin altistuksen suhteen toisistaan poikkeavia ryhmiä ja sen jälkeen kerätään tietoa sairauden esiintyvyydestä eri ryhmissä. Tekemällä altistusmittaukset ennen sairauden ilmaantumista voidaan varmistua yhteyden ajallisesta suunnasta (altistus ennen vastetta). Altistuksen ja sairauden välistä yhteyttä mitataan yleensä esiintyvyyssuhteella eli suhteellisella riskillä (rate ratio, relative risk, RR), joka on intuitiivisesti helpommin ymmärrettävissä kuin ristitulosuhde. Vaihtoehtoisesti voidaan käyttää esiintyvyyseroa (rate difference). Katso taulukko 6.3.



	Sairastuneet	Henkilövuodet
Altistuneet	A	B
Altistumattomat	C	D

**TAULUKKO 6.3 Kohorttitutkimuksen analyysi**

Taulukossa A kuvaa sairastuneiden määrää altistuneiden joukossa, B henkilövuosien määrää altistuneilla, C sairastuneiden määrää altistumattomilla ja D henkilövuosien määrää altistumattomien joukossa. Esiintyvyyssuhde kuvaa sairauden yleisyyttä tutkimusryhmässä suhteessa sen esiintyvyyteen vertailuryhmässä. Esiintyvyyssuhde,  $RR = \lambda_1/\lambda_0 = a/b : c/d$ . Esiintyvyysero,  $RD = \lambda_1 - \lambda_0 = a/b - c/d$

Kohorttitutkimuksella voidaan selvittää saman altistuksen vaikutusta moneen eri sairauteen, ja se on tehokkain tapa tutkia harvinaisen altistuksen vaikutuksia. Altistuksen vaikutus ilmaistaan esiintyvyyssuhteena tai esiintyvyyserona. Jos merkitään sairauden esiintyvyyttä altistumattomassa ryhmässä  $\lambda_0$ :lla ja altistuneessa ryhmässä  $\lambda_1$ :llä, niin esiintyvyyssuhde  $RR = \lambda_1/\lambda_0$  ja esiintyvyysero  $RD = \lambda_1 - \lambda_0$ . Esiintyvyyssuhde kuvaa suhteellista riskiä ja esiintyvyysero absoluuttista riskiä. Kohorttitutkimuksen tilastolliseen voimaan vaikuttaa sairaustapausten määrä eri altistusryhmissä. Mitä yleisempi tutkittava sairaus tutkimusväestössä on, sitä suurempi tilastollinen voima. Myös altistusjakauma on hyvin tärkeä. Yleensä tilastollinen tehokkuus on suurin, kun vertailtavissa ryhmissä on sama määrä sairaustapauksia. Kohorttitutkimuksessa vertailtavat ryhmät muodostetaan yleensä seurannan alkuhetkellä. Koska altistukset muuttuvat usein ajan myötä, on tärkeää kerätä tietoa altistuksesta eri ajankohtina. Kohorttitutkimuksessa on tärkeää varmistua altistusryhmien vertailukelpoisuudesta, sillä yleensä altistukseen liittyy muita valikoivia tekijöitä. Muun muassa altistuneiden koulutus, tulotaso ja elintavat saattavat poiketa altistumattoimien vastaavista. Tämän vuoksi sopivan vertailuryhmän löytäminen on keskeisen tärkeää. Seurannan kattavuus, eli kaikkien sairaustapausten identifiointi, on kohorttitutkimuksissa erittäin tärkeää. Jotta vaikutuksia voitaisiin tutkia, altistuksen alusta täytyy olla kulunut riittävän pitkä aika. Syövän kohdalla tämä on usein kymmenen vuotta tai enemmän. Sairaustapaukset täytyy todeta seurannassa samoilla kriteereillä ja yhtä kattavasti vertailtavissa ryhmissä, jotta vältetään informaatioharha. Joissain tapauksissa altistuminen saattaa vaikuttaa sairauden toteamiseen. Esimerkiksi altistuneelle ammattiryhmälle saatetaan tehdä terveystarkastuksia todetun tai epäillyn haitan vuoksi, mikä voi johtaa sairaustapausten kattavampaan toteamiseen kuin muissa ryhmissä.

Satunnaistetun kokeen vahvuuksia ovat se, että tutkija voi määrittellä altistuksen ja muodostaa tutkittavat ryhmät satunnaisesti maksimoiden vertailukelpoisuuden. Satunnaistetun kokeen analyysi muistuttaa kohorttitutkimuksen analyysiä. Koska tutkimushenkilöiden tahallinen altistaminen haitallisille tekijöille ei ole eettisesti hyväksyttävää, satunnaistettu koe ei kuitenkaan yleensä sovellu säteilyepidemiologiaan. Poikkeuksena ovat joidenkin säteilyaltistuksen sisältävien kokeiden, kuten sädehoitoa tai röntgentutkimuksia koskevien tutkimusten analyysi haittavaikutusten osalta. Näitä kokeita ei kuitenkaan ole suunniteltu ensisijaisesti säteilyn haittavaikutusten tutkimusta varten.

### 6.3 Riskimallit

Suhteellisen riskin malli kuvaa tutkittavan tekijän  $x$  vaikutusta suhteessa muiden riskitekijöiden vaikutukseen eli taustaesiintyvyyteen  $\lambda_0$ . Tällöin eri tekijöiden ajatellaan vaikuttavan samaan prosessiin, eikä niiden tarvitse olla yksinään riittäviä vaikutuksen ilmenemiseen. Jos sairauden esiintyvyyttä altistuneessa väestössä merkitään  $\lambda_1$ :lla,  $x$  kuvaa altistumista tutkitavalle tekijälle, ja  $\beta$  on riskikerroin joka kuvaa  $x$ :n vaikutuksen suuruutta, niin voidaan merkitä:

$$\text{Suhteellisen riskin malli: } \lambda_1 = \lambda_0 + \beta x \lambda_0 = \lambda_0 (1 + \beta x)$$

Vaikutusta kuvataan tällöin esiintyvyyssuhteella eli riskisuhteella (rate ratio,  $RR = \lambda_1 / \lambda_0$ ). Jos vaikutusta ei ole ( $\lambda_1 = \lambda_0$ ), esiintyvyyssuhde saa arvon  $RR = 1$ . Suojavaikutusta vastaa arvo  $RR < 1$ . Tutkittavan tekijän vaikutusta voidaan kuvata myös suhteellisella lisäriskillä (excess relative risk, ERR), joka muodostuu yllä olevan yhtälön loppuosasta ( $\beta x \lambda_0$ ). Lisäriski lasketaan vähentämällä esiintyvyyssuhteen arvosta 1 ( $ERR = RR - 1$ ). Lisäriski siis kuvaa esiintyvyyden muutosta  $\lambda_0$ :aan lisättävän kerrannaisen avulla siten, että vaikutuksen puuttumista vastaa arvo  $ERR = 0$  ja suoja-vaikutusta negatiivinen arvo ( $ERR < 0$ ).

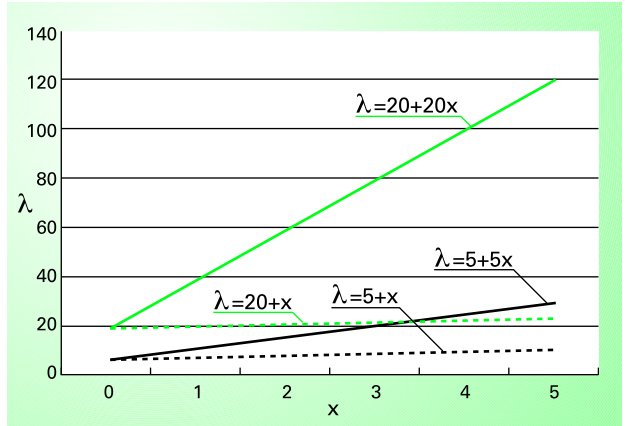
Absoluuttisen riskin malli kuvaa tutkittavan tekijän  $x$  vaikutusta riippumatta muista siihen vaikuttavista tekijöistä. Tämä vastaa vaikutusta, jossa  $x$  on vastetapahtuman riittävä syy.

$$\text{Absoluuttisen riskin malli: } \lambda_1 = \lambda_0 + \beta x$$

Absoluuttisen riskin mallissa vaikutusta kuvataan esiintyvyyserolla tai riskierolla (rate difference, RD, tai excess absolute risk, EAR). Jos vaikutusta

ei ole ( $\lambda_1 = \lambda_0$ ), riskiero saa arvon nolla. Suojavaikutusta vastaa negatiivinen arvo ( $RD < 0$ ).

Suhteellisen ja absoluuttisen riskin malleja voidaan havainnollistaa esimerkiksi, jossa kuvataan samansuuruisen altistuksen  $x$  vaikutusta kahdessa väestössä, jotka poikkeavat toisistaan sairauden yleisyyden suhteen. Toisessa taustailmaantuvuus  $\lambda_0$  on 5 ja toisessa 20, riskikerroin  $\beta$  saa molemmissa väestöissä ja riskimalleissa arvon 1 (kuva 6.1).

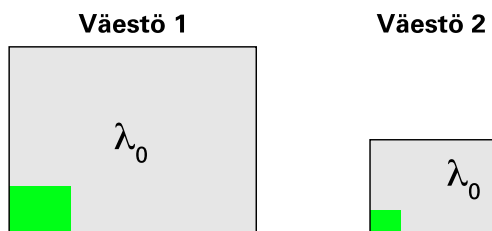


**KUVA 6.1 Absoluuttinen lisäriski ja suhteellinen lisäriski**  
 $x$  = altistus,  $\lambda$  = ilmaantuvuus

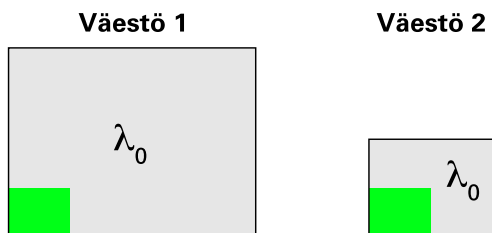
Eri riskimalleista olkoon esimerkkinä tupakoinnin määrän yhteys sepelvaltimotautiin. Suhteellisen riskin mallia käyttäen tupakoinnin vaikutus sepelvaltimotaudin ilmaantuvuuteen voidaan kuvata yhtälöllä  $\lambda_1 = \lambda_0 + \beta x \lambda_0$ , jossa  $\lambda_0$  on ilmaantuvuus tupakoimattomilla,  $x$  on poltetun tupakan määrä savukkeina päivässä ja  $\beta$  on riskikerroin, joka ilmaisee kuinka paljon riski suurenee tupakointiyksikköä kohti. Jos ilmaantuvuus tupakoimattomilla on 1/1000 ja  $\beta$  saa arvon 0,005, on sepelvaltimotaudin vaara kymmenen savuketta päivässä polttavilla  $\lambda_1 = \lambda_0 + \beta x \lambda_0 = 0,001 + 0,005 \cdot 10 \cdot 0,001 = 0,00105$ . Ilmaantuvuussuhde on siis  $\lambda_1/\lambda_0 = 0,00105/0,001 = 1,05$  ja suhteellinen lisäriski  $1,05 - 1 = 0,05$ . Mikäli suhteellisen riskin malli on oikea, tupakoinnin vaikutus on suhteessa sepelvaltimotaudin esiintymiseen ja ilmaantuvuussuhde on vakio eri väestöissä. Näin ollen suhteellinen vaikutus eli osuus tapauksista on vakio, mutta absoluuttinen vaikutus eli tapausmäärä suhteessa väestön kokoon on suurempi sellaisessa väestössä, jossa sairautta esiintyy muutenkin enemmän ( $\lambda_0$  on suuri).

Jos taas tupakoinnin ja sepelvaltimotaudin yhteys kuvataan absoluuttisen riskin mallia käyttäen, sepelvaltimotaudin ilmaantuvuus suhteessa tupakoinnin määrään ilmaistaan samoja merkintöjä käyttäen yhtälöllä  $\lambda_1 = \lambda_0 + \beta x$ . Jos riskikerroimen arvoksi  $\beta$  saadaan 0,005 ja ilmaantuvuus vertailuryhmässä on 1/1000, on sepelvaltimotaudin vaara 10 savuketta päivässä polttavilla  $\lambda_1 = \lambda_0 + \beta x = 0,001 + 0,005 \cdot 10 = 0,051$ . Ilmaantuvuusero on  $\lambda_1 - \lambda_0 = 0,051 - 0,001 = 0,05$ . Mikäli

absoluuttisen riskin malli on oikea, tupakoinnin vaikutus on riippumaton sepelvaltimotaudin esiintymisestä ja riskiero on vakio eri väestöissä. Näin ollen absoluuttinen vaikutus eli tapausmäärä suhteessa väestön kokoon on vakio eri väestöissä. Suhteellinen vaikutus eli osuus tapauksista on kuitenkin suurempi sellaisissa väestöissä, joissa taustailmaantuvuus ( $\lambda_0$ ) on pieni (katso kuva 6.2).



Suhteellinen riski on vakio väestöissä 1 ja 2.



Absoluuttinen riski on vakio väestöissä 1 ja 2.

**KUVA 6.2** Suhteellinen riski ja absoluuttinen riskimalli. Vaalea laatikko edustaa taustaesiintyvyyttä ( $I_0$ ) ja varjostettu alue altistuksen vaikutusta.

Samoja malleja voidaan soveltaa kun tutkitaan useamman tekijän yhteisvaikutusta. Eri tekijöiden ( $x_1$  ja  $x_2$ ) vaikutus voi olla kerrannainen (multiplikatiivinen) tai additiivinen. Suhteellisen riskin mallissa eri tyyppiset yhteisvaikutukset voidaan merkitä:

Multiplikatiivinen yhteisvaikutus: 
$$\lambda_1 = \lambda_0 \cdot \beta_1 x_1 \cdot \beta_2 x_2$$

Additiivinen yhteisvaikutus: 
$$\lambda_1 = \lambda_0 \cdot \beta_1 x_1 + \lambda_0 \cdot \beta_2 x_2$$

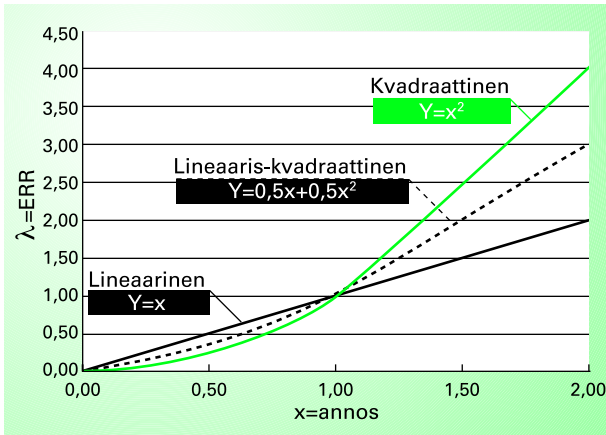
Absoluuttisen riskin mallissa vastaavasti:

Multiplikatiivinen yhteisvaikutus: 
$$\lambda_1 = \lambda_0 + \beta_1 x_1 \cdot \beta_2 x_2$$

Additiivinen yhteisvaikutus: 
$$\lambda_1 = \lambda_0 + \beta_1 x_1 + \beta_2 x_2$$

## 6.4 Annosvaste

Altistuksen vaikutusta sairauden esiintymiseen voidaan kuvata annoksen funktiona. Annos-vastemalli on erityisen tärkeä sovellettaessa suurista annoksista johdettuja riskiarvioita pienemmille annoksille. Ilman kynnyksarvoa esiintyviä vaikutuksia voidaan kuvata yksinkertaisilla annos-vaste malleilla. Yleisimmin käytetty on lineaarinen malli, jossa vaikutus lisääntyy yhtä paljon annosyksikköä kohti eri annostasoilla (kuva 6.3).



KUVA 6.3 Lineaarinen, kvadraattinen ja lineaaris-kvadraattinen annosvaste

**Lineaarinen annosvaste:**  $I = a + bx$  (kuvan esimerkissä  $I = x$ )

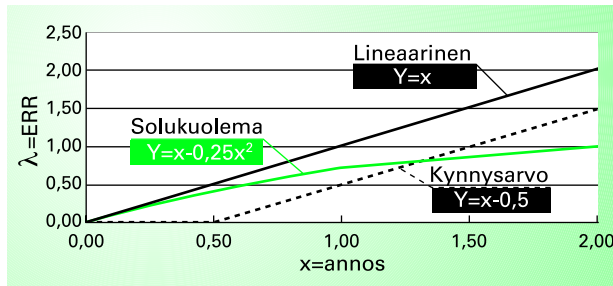
**Kvadraattinen annosvaste:**  $\lambda = \alpha + \beta x^2$  (kuvan esimerkissä  $\lambda = x^2$ )

**Lineaaris-kvadraattinen annosvaste:**  $\lambda = \alpha + \beta x + \gamma x^2$

Edellä kuvatussa esimerkissä käytettiin lineaarista mallia, eli riskin lisäksi poltettua savuketta kohti oli vakio. Kvadraattisessa mallissa annoksen neliöön liittyvää vaikutusta kuvaava neliötermi voi olla negatiivinen, jolloin vaikutus annosyksikköä kohti heikkenee suurilla annoksilla, tai positiivinen, jolloin vaikutus annosyksikköä kohti suurenee isoilla annoksilla, kuten kuvassa 6.3. Lineaaris-kvadraattinen malli sisältää sekä lineaarisen että kvadraattisen komponentin. Tästä esimerkkinä on lineaariskvadraattisessa mallissa suurilla säteilyannoksilla tapahtuvaa solu-kuolemaa kuvaava negatiivinen kvadraattinen annostermi (kuva 6.4). Lineaarisen ja kvadraattisen komponentin suhdetta kuvataan näiden kahden kertoimen suhteella,  $t = \gamma / \beta$ , jossa  $t$ :n suuruus kuvaa poikkeamaa lineaarisuudesta eli kvadraattisen komponentin suuruutta siten, että arvo  $t = 0$  vastaa puhdasta lineaarista mallia ja ääretön puhdasta kvadraattista mallia. Riskin ekstrapolaatio suurilta annoksilta pienille voidaan kuvata

korjaustekijällä  $k$ , joka saadaan kvadraattisen kertoimen suhteena lineaariseen kertoimeen:  $\beta x + \gamma x^2 = \delta (x + kx^2)$ .

Joissain tapauksissa voi esiintyä kynnyisarvo, jonka alapuolella vaikutuksia ei ilmene (kuva 6.4).



**KUVA 6.4**

Lineaarinen vaste ilman kynnyisarvoa, lineaarinen annosvaste, jossa kynnyisarvo, sekä lineaaris-kvadraattinen annosvaste, jossa on negatiivinen kvadraattinen termi.

Eri annos-vastemallien vertailu on kuitenkin käytännössä vaikeaa, sillä pienillä annoksilla mallien tulokset eroavat toisistaan vain hieman ja niiden luotettava erottelu vaatii huomattavan suuren aineiston, jossa on laaja annosjakauma. Lisäksi annos-vaste saattaa riippua muun muassa syöpätyypistä, annoksesta, annosnopeudesta, altistusikästä tai seurannan pituudesta. Empiirisestä aineistosta saadut tulokset riippuvat näiden tekijöiden jakaumasta käytetyssä aineistossa. Annosvastemalleja voi myös soveltaa eri riskimalleihin. On esimerkiksi täysin mahdollista, että samassa aineistossa suhteellinen riski noudattaa lineaaris-kvadraattista mallia ja absoluuttinen riski lineaarisesta mallia.

Mittausvirhe vaikuttaa annosvasteeseen, vaikka ei aiheuttaisikaan harhaa annosarvioihin. Mittausvirhe pienentää riskikerrointa ja jos virhe on suhteessa annokseen, riskikerroin pienenee eniten suurilla annoksilla.

## KIRJALLISUUTTA

Hernberg S. Epidemiologia ja työterveys. Helsinki: Työterveyslaitos, 1998.

Gordis L. Epidemiology. Philadelphia: WB Saunders, 2000.