



OULUN YLIOPISTO  
UNIVERSITY of OULU

# *Case-control* – Pesästetyt ja muut vastepainotteiset tapaus-alustaotosasetelmat

Esa Läärä

Suomen epidemiologian seuran seminaari  
Helsinki 22.4.2009

Suomen epidemiologian seura ja Duodecim

# “Case-control studies” (CC)

## Synonyymejä

- ▶ case-comparison s., case-compeer s.
- ▶ case-referent s., case-base s.

## Muunnelmia

- ▶ traditional “case-noncase” s., epidemic CC s., cumulative incidence CC,
- ▶ (incidence) density CC, nested CC,
- ▶ case-cohort s., case-base s.,
- ▶ case-crossover s., case-time-control s.,
- ▶ case-case s.,
- ▶ case-specular s., . . .



# “Case-control study” joillakin kielillä

- ▶ juhtkontrolluuring,
- ▶ fall-kontrollstudie,
- ▶ case-kontroll studie, case-kontrol-studie,
- ▶ studio caso-controllo, estudio de casos y controles,
- ▶ **étude cas témoin**  
*témoin* = 'silminnäkijä', 'todistaja', 'sekundantti'



# “Case-control study” suomeksi

*Lääketieteen englanti-suomi-sanakirja:*

- ▶ tapaus-verrokkitutkimus,
- ▶ verrokkiparitutkimus,
- ▶ sairauslähtöinen tutkimus.

*Lääketieteen termit:*

“epidemiologinen tutkimus, jossa verrataan toisiinsa esim. kutakin sairastunutta (“tapaus”) ja hänelle pareiksi valittuja yhtä tai useampaa muuten vastaavaa mutta terveenä pysynyttä henkilöä (verrokki)”

- ▶ *Ovatko verrokki aina tapauksille paritettuja?*
- ▶ *Minkä suhteen verrataan ja missä tarkoituksessa?*



## Esim. Näytönastekatsaus

- ▶ “Amerikkalaisessa retrospektiivisessä tutkimuksessa seurattiin naisia, joille vuosina 1960 - 1993 oli tehty ennaltaehkäisevä molempien rintojen poisto.
- ▶ Kyseessä oli tapaus-verrokki -tutkimus.
- ▶ Verrokkeina [ . . . ] sisaruksia, joille ei [ . . . ] rintojen poistoa.
- ▶ Hoidetuista potilaista 214 kuului korkean riskin ryhmään ja seuranta aikana [!] heistä kolmelle (1.4 %) kehittyi rintasyöpä.
- ▶ Verrokeita oli 403 ja heistä 156:lle (38.7 %) kehittyi rintasyöpä, joten rintasyövän esiintyvyys hoidetuilla potilailla väheni vähintään 90 %.”



## Esim. Näytönastekatsaus

- ▶ “Kliininen tapaus-verrokkitutkimus, jossa
- ▶ selvitettiin 217 suomalaisen yliopisto-opiskelijan [. . .] osittain puhjenneiden alaviisaudenhampaiden määrän vaikutusta syljen kariesta aiheuttaviin bakteerimääriin.
- ▶ Tuloksissa todettiin tilastollisesti merkitsevästi enemmän korkeita laktobasilli- [. . .] ja *S. mutans*-määriä [. . .] opiskelijoilla, joilla oli 1-3 viisaudenhammasta osittain puhjenneina verrattuna tapauksiin, joilla ei ollut tällaisia hampaita.”

Artikkeli nak06460 ©2009 Kustannus Oy Duodecim



## Esim. Näytönastekatsaus

- ▶ “[...] retrospektiivisessä tapaus-verrokkitutkimuksessa arvioitiin potilaan iän vaikutusta tulehduksellisen sairauden vuoksi leikattujen potilaiden ennusteeseen.
- ▶ Tutkimukseen otettiin 30 yli 60-vuotiasta potilasta, jotka leikattiin saman kirurgin toimesta 10 vuoden kuluessa.
- ▶ Verrokkiryhmän muodosti 75 sukupuolen, leikkausajan-kohdan ja -menetelmän perusteella samaistettua [!] alle 60-vuotiasta.
- ▶ Vanhemmassa ikäryhmässä komplikaatioiden määrä oli suurempi [...] ja leikkauksen [...] ja sairaalahoidon kesto [...] olivat pitemmät myös sen jälkeen, kun vakioitiin [...]”



## Esim. Erikoislääkäarin uutiset

- ▶ “[...] väitöskirjassa osoitettiin, että keskimääräisesti ottaen psykodynaamiseen terapiaan hakeutuneiden oireet vähenevät ja persoonallisuuden rakenne ja ihmissuhteet kohenevat hoidon aikana.
- ▶ Kyseessä on tapaus-verrokkitutkimus, jossa
- ▶ hoidettuja verrattiin sellaisiin potilaisiin, jotka hakeutuivat terapiaan mutta eivät voineetkaan aloittaa hoitoa syystä tai toisesta.
- ▶ Terapiat olivat pitkiä, ja hoidettuja potilaita oli 55.”

*Duodecim* 2002; 118(24): 2513.



## Kliininen “tapaus-verrokkitutkimus”?

- ▶ Melkein mikä tahansa kliininen tutkimus sisältää “tapauksia”, joita verrataan joihinkin “verrokkeihin”.
- ▶ Silti mikään edellä ei ole “tapaus-verrokkitutkimus” – siten kuin epidemiologiassa ymmärretään – vaan “kohorttitutkimus” kohteena ao. potilasjoukon ennuste!
- ▶ DoE5: *“It is not correct to call ‘case-control study’ any comparison of a group of people having a specific outcome with another group free of that outcome”*

### *Mistä lähteistä katsausten kirjoittajat ovat ammentaneet oppinsa?*

Eivät ole ainakaan huolella lukeneet Duodecimin julkaisemia oppimateriaaleja.



# Duodecimin oppi- ja käsikirjat

*Käypä hoito -käsikirja* (2008):

- ▶ “**Tapaus-verrokkitutkimus (retrospektiivinen asetelma – analyysi taudista altistukseen):** Ryhmät tunnistetaan tuloksen perusteella ja altistusta etsitään taannehtivasti.”

Uhari & Nieminen (2001). *Epidemiologia ja biostatistiikka:*

- ▶ “*Tapaus-verrokkitutkimus*  
Tutkija valitsee tapaukset ja niille vertailtavat kontrollit tutkittavan taudin mukaan (sairaat tulevat tapauksiksi ja terveet kontrolleiksi). Aiempaa altistumista verrataan näissä kahdessa ryhmässä.” (s. 35).



# Näyte ulkomaisista oppikirjoista

Gordis, L. (2008). *Epidemiology, 4th Ed.*  
Elsevier-Saunders, Philadelphia. s. 177:

- ▶ “design of a *case-control study* [...] we identify a group of individuals with that disease (called cases) and,
- ▶ for the purposes of comparison, a group of people without that disease (called controls).
- ▶ We determine what proportion of the cases [ and ] controls were exposed and what proportion were not.”

***Edustava esimerkki alan opetuksesta 2000-luvulla?***



## Ero “kohorttitutkimuksiin”?

- ▶ “What distinguishes the two study designs is whether the study begins with diseased and non-diseased people [...] (case-control study) or with exposed and nonexposed people (cohort study).” (Gordis 2008, s. 179)
- ▶ “case-control studies involve looking backward from disease to exposure, in contrast to cohort studies, which look forward from exposure to disease.”  
Fletcher & Fletcher (2005). *Clinical Epidemiology*, s. 92

*Näinkö on?*



## “Backward” – “begins with cases”?

- ▶ Tapaus-verrokkitutkimuksella “takeneva suuntaus”?  
seuraus → syyt ; vaste → altistus
- ▶ Onko tutkimuksen logiikka täysin päinvastainen kuin  
“kohorttitutkimuksen” joka on “etenevä”?  
syyt → seuraus

**Esim.** Tartuntatautiepidemioiden selvittäminen

- ▶ aina “tapauslähtöistä” (*begins with cases*),
- ▶ silti asetelmana käytetään usein “kohorttitutkimusta”.  
(Nuorti & Kuusi. *Duodecim* 2002; 118: 2038–45)

*Missä siis ero?*



# “Case-control studies” – tosiollemus

- ▶ “In the 1970s [...] case-control studies [...] viewed as *efficient sampling schemes* of the disease experience of the underlying [...] *cohorts*”  
(MacMahon & Trichopoulos 1996. *Epidemiology*)
- ▶ “[...] the modern conception of the case-control study as involving *sampling*, on the basis of outcome, from an ongoing real or imagined cohort study [...]”  
(Breslow 2005, kirjassa *Handbook of Epidemiology*)



# “Case-control studies” – tosiollemus

Eräitä virstanpylväitä

- ▶ Mantel (1973): “Synthetic retrospective studies”,
- ▶ Kupper ym. (1975): “Hybrid epidemiological design”,
- ▶ Miettinen (1976): “Case-referent studies”,
- ▶ Thomas (1977),
- ▶ Prentice & Breslow (1978),
- ▶ Miettinen (1982, 1985): “Study base, Case-base”,
- ▶ Prentice (1986): “Case-cohort”,
- ▶ Wacholder (1996): “Data missing by design”.



## “Case-control studies” – tosiolemus

*Yhteistä* “kohortti-” ja “tapaus-verrokkitutkimuksille”?

- ▶ tutkimuskysymys: missä määrin altistuminen tekijöille  $A_1, A_2, \dots$  vaikuttaa vasteen  $B$  esiintyvyyteen.
- ▶ tapausten lähde: määrätty **tutkimusväestö** tietyssä **aikana** → **tutkimusalusta** t. -**pohja** (*study base*, viroksi *uuringualus*).

*Ero?* – Altistumistietojen hankinta (Miettinen 1985)

- ▶ “kohorttitutkimus” – **kokonaistutkimus** (*census*): koko tutkimusväestö/-alusta ,
- ▶ “tapaus-verrokkitutkimus” – **otanta** (*sampling*):
  - (1) kaikki tapaukset (tai valtaosa) sekä
  - (2) tutkimusväestöä/-alustaa edustava otos.



## Yksi suomalainen oppikirja!

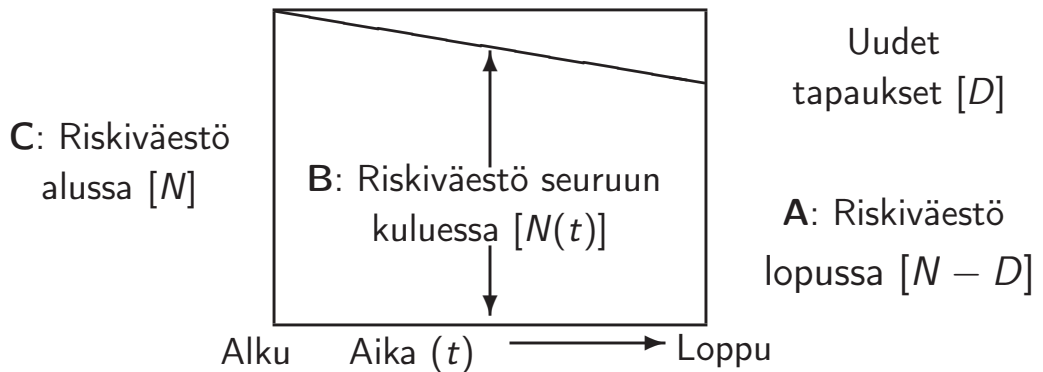
- ▶ “Otokseen perustuvissa tutkimuksissa väestöpohjasta valitaan edelleen kaikki kyseisen sairauden tapaukset, mutta vain *otos* tutkimuspohjasta.
- ▶ Tätä tutkimustyyppiä on perinteisesti kutsuttu *tapauserrokitutkimukseksi* (case-referent study, case-control study, [. . .] toisinaan retrospective study).
- ▶ vain murto-osalta kaikista tutkimuspohjan jäsenistä pitää hankkia altistumistiedot, mikä säästää työtä ja varoja.
- ▶ Luotettavan tuloksen saaminen edellyttää, että verrokkitodellakin edustavat tutkimuspohjaa [. . .] verrokkit voidaan valita usealla tavalla.

Hernberg, S. (1998). *Epidemiologia ja työterveys*, 2. p.



## Ideaallinen pitkittäinen tutkimusalusta

- ▶ suljettu tutkimusväestö (kohortti) riskialttiita yksilöitä,
- ▶ seuruu kiinteän riskijakson yli,
- ▶ ei sensurointeja eikä kilpailevia tapahtumia.



Vaihtoehtoiset otantakehikot: **A**, **B** ja **C**

## Kokonaistutkimus – “kohorttitutkimus”

Kaksiarvoinen altistus: 1 = “kyllä”; 0 = “ei”.

Karkea analyysi tiivistyy  $2 \times 2$ -taulukkoon:

|             | altistunut      | altistumaton    | kaikki    |
|-------------|-----------------|-----------------|-----------|
| tapaukset   | $D_1$           | $D_0$           | $D$       |
| ryhmän koko | $N_1$           | $N_0$           | $N$       |
| henkilöaika | $Y_1$           | $Y_0$           | $Y$       |
| ilm. osuus  | $Q_1 = D_1/N_1$ | $Q_0 = D_0/N_0$ | $Q = D/N$ |
| ilm. tiheys | $I_1 = D_1/Y_1$ | $I_0 = D_0/Y_0$ | $I = D/Y$ |

Voidaan laskea mm. ilmaantumistiheyksien suhde

$$IR = \frac{D_1/Y_1}{D_0/Y_0} = \frac{D_1/D_0}{Y_1/Y_0},$$

= riskitiheyksien suhteen (*hazard ratio*, HR) estimaatti.



## Otanta – vaihtoehtoiset asetelmat

### A: Perinteinen otanta (“*case-noncase*” sampling):

- ▶ Otos poimitaan niistä  $N - D$  yksilöstä, jotka vielä riskialttiita (“terveitä”) seuruun lopussa

### B: Tiheysotanta (*density sampling*):

- ▶ Otos poimitaan tutkimusalueesta t. riskiväestöstä seuruun kuluessa
- ▶ “*Nested case-control design*”-variantti (NCC)  
Pieni otos (1–5 “verrokkia”) poimitaan riskiväestöstä kunkin uuden tapauksen diagnosihetkenä t

### C: Tapaus-kohorttiotanta (*case-cohort sampling* CC):

- ▶ Otos eli **alikoortti** (*subcohort*) poimitaan koko kohortista eli riskiväestöstä ( $N$ ) seuruun alussa.



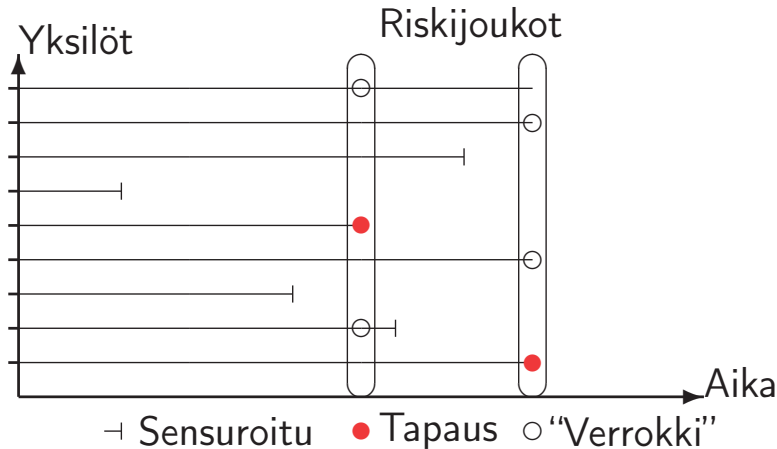
## “Nested case-control (NCC)”

- ▶ Tiheysotannan variantti: **aikakaltaistettu otanta riskijoukosta** (*time-matched risk-set sampling*)  
Kunakin hetkenä  $t$ , kun uusi tapaus diagnosoidaan, ryhmä verrokkeja poimitaan sillä hetkellä **riskijoukkoon** (*risk-set*) kuuluvista tutkimusväestön jäsenistä.
- ▶ Joissakin teksteissä (mm. Silva 1999, Rothman et al. 2008) NCC on mikä tahansa “tapaus-verrokkitutkimus”, jossa tapausten lähde on hyvin määritelty väestö.
- ▶ *Nested* suomeksi?
  - (kohortin) sisäinen
  - upotettu (Hernberg 1998)
  - pesästetty, pesitetty? (viroksi *pesastatud*)



## NCC – *étude cas-témoins nichée*

Aina kun uusi tapaus diagnosoidaan, pieni ryhmä “verrokkeja” (tässä 2/tapaus) poimitaan riskijoukosta.



NB. “Verrokiksi” valitusta voi myöhemmin tulla myös tapaus.

# NCC – Riskijoukko-otanta

## Riskijoukko

- = tapauskohtainen otantakehikko,
- = tutkimusväestön ne yksilöt ( $\times$ ), jotka ovat alttiina vasteen saamiselle tapauksen dg-hetkellä.

Tapaus

Terve loppuun asti

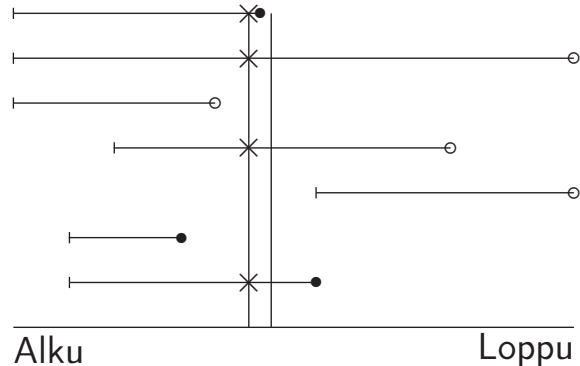
Varhainen kato

Myöhempi tulo

Vielä myöhempimpi

Varhempi tapaus

Myöhempi tapaus



# Eri asetelmien käyttö

## A oppikirja-asetelma, “tapaus/ei-tapaus”-otanta

- ▶ akuutit epidemiat,
- ▶ perinataalivasteet (jolloin tutkimusalusta poikittainen!),
- ▶ syöpä, sydäntaudit, jne? – tuskin koskaan,

## B tiheysotanta, useimmiten NCC

- ▶ ylivoimaisesti käytetyin kroonisissa taudeissa,
- ▶ ainoa looginen asetelma avoimessa väestössä,

## C tapaus-kohorttiotanta

- ▶ monia vasteita tutkittavissa samalla otoksella (vrt. “Tapaus-verrokkitutkimuksessa voidaan tutkia vain yhtä valittua sairautta” U & N 2001, s. 35!),
- ▶ suosio kasvaa erit. geneettisessä epidemiologiassa.

***Monet oppikirjat eivät esittele lainkaan asetelmia B ja C, vaikka juuri niitä käytetään kroonisissa taudeissa!***



# “Suhteellisen riskin” estimointi

Kaksiarvoinen altistus: 1 = “kyllä”; 0 = “ei”.

Karkea analyysi tiivistyy  $2 \times 2$ -taulukkaan:

|                   | altistunut | altistumaton | kaikki |
|-------------------|------------|--------------|--------|
| tapaukset         | $D_1$      | $D_0$        | $D$    |
| otos (“verrokki”) | $C_1$      | $C_0$        | $C$    |

**Altistumisen jakosuhte** (*exposure odds ratio* EOR):

$$\begin{aligned} \text{EOR} &= \frac{D_1/D_0}{C_1/C_0} \\ &= \frac{\text{tapaukset: altistuneet/ei-altistuneet}}{\text{otosyksilöt: altistuneet/ei-altistuneet}} \end{aligned}$$



# “Suhteellisen riskin” estimointi

Tiheysotannassa **B**

altistumisjako  $C_1/C_0$  otoksessa (“verrokeilla”)

estimoit harhattomasti

altistumisjakoa  $Y_1/Y_0$  koko alustassa,

jossa  $Y_1$  ja  $Y_0$  ovat henkilövuosien määrät riskitekijälle altistuneiden ja altistumattomien ryhmässä koko siinä tutkimusväestössä/-alustassa, josta tapaukset ovat peräisin.



## “Suhteellisen riskin” estimointi

- ⇒ EOR tapausten ja “verrokkien” välillä on harhaton ja tehokas estimaattori riskitiheyssuhteelle (HR) altistusryhmien välillä ilman “harvinaisen taudin oletusta”.
- ▶ Vastaavassa “kohorttitutkimuksessa” HR:n estimaatti on siis ilmaantumistiheyksien suhde

$$IR = \frac{D_1/Y_1}{D_0/Y_0} = \frac{D_1/D_0}{Y_1/Y_0} \approx \frac{D_1/D_0}{C_1/C_0} = EOR$$

### “Verrokkien” rooli?

- Ei ole edustaa terveinä pysyviä “ei-tapauksia”, vaan
- + antaa informaatiota altistuksen jakaumasta koko tutkimusväestössä/-alustassa



# Terminologian osuvuus?

- ▶ Otosta riskiväestöstä tai tutkimusalustasta kutsutaan “verrokkiryhmäksi” (“control group”)
  - ▶ Ajatuksena takaperoinen vertailu “tapausten” (sairaat) ja “ei-tapausten” (terveet) välillä.
  - ▶ Oikea vertailun kohde: vasteen esiintyvyys, riski t. riskitiheys altistuneiden ja altistumattomien välillä.
  - ▶ “Verrokki” – myös “altistumaton” vertailuryhmäläinen!
- ⇒ “Verrokki” ja “tapaus-verrokki-” (“control, case-control”) harhaanjohtavia termejä kun puheena on otos t. otanta tutkimusalustasta – hylättäviä!

***Mitä tilalle?***



## “Case-control study” toisin sanoin?

- ▶ case-referent design, case-base sampling (Miettinen)
- ▶ outcome-dependent, outcome-based, outcome-selective,
- ▶ response-dependent, response-based, response-selective,
- ▶ choice-based sampling (erityisesti mikroekonometriassa).

“*Control*” → *samplee?* (otosyksilö)

Suomeksi?

- ▶ vastepainotteinen otanta,
- ▶ tapaus-riskiväestötanta, tapaus-alusta-asetelma, tapaus-alustaotostutkimus, ...?

“*Verrokki*” → otokki, otonen, otoke, ote, ...?



# Pohdinta

- ▶ Tapaus-alusta-asetelmien tosiolemus tiedetty jo  $> 30$  v.
- ▶ Silti lukuisat oppikirjat ja kurssit ympäri maailmaa jatkavat harhaopin levittämistä “tapaus-verrokki”-tutkimuksista vielä 2010 jälkeen.
- ▶ Vakiintuneet mutta huonosti sopivat sanat kuten “*control*” (“verrokki”) ylläpitävät vääriä käsityksiä.
- ▶ Vanhan epidemiologin toive:  
Ainakin Suomen epidemiologian tutkijat ja tuntijat opettaisivat sanoista huolimatta asetelmat oikein.

***Sanastotyö siis jatkuu: sanat ja opit kohdalleen!***



# Kiitokset!

- ▶ SES, Duodecim, FinOHTA, Suomen Syöpärekisteri,
- ▶ sanastoseminaarien 22.3.2006 ja 24.11.2008 alustajat ja osallistujat,
- ▶ suomen kielen asiantuntijat:  
Pirjo Hiidenmaa ja Harri Mantila,
- ▶ kommentoijat:  
mm. Katriina Heikkilä (Bristol) ja Lyly Teppo,
- ▶ tekninen toimittaja: Eeva Vaaramo,
- ▶ sanastotoimikunnan aktiivit: Tapio Luostarinen (siht.), Anna-Maija Virtala, Outi Lyytikäinen, Seppo Sarna, Hilka Riihimäki, Timo Hakulinen, Matti Hakama (pj.)



## Epilogi: Vuoden epidemiologi – riskitekijät?

- ▶ mies, tilastotieteilijä, Helsingistä, Jaakko,...  
AST-syndrooma?
- ▶ Tapaukset: Vuoden epidemiologit 2004-9;  $D = 6$
- ▶ “Verrokki”: SES:n muut jäsenet 22.4.2009;  $C = 144$

|          | AST <sup>+</sup> | AST <sup>-</sup> | kaikki |
|----------|------------------|------------------|--------|
| tapaus   | 4                | 2                | 6      |
| verrokki | 4                | 140              | 144    |

- ▶ Vallitsevuuksien vetosuhte = altistumisen jakosuhte

$$\text{EOR} = \frac{4/2}{4/140} = 70 \quad [95\% \text{ CI: } 9.8 - 500]$$

$$P < 0.00005 \quad \text{*****} \quad [\text{Fisherin “eksakti”}]$$

- ▶ AST = Aktiivinen sanastotoimikuntalaisuus.

