



# Matching i kohortstudier

---

## MEDISIN OG TALL

ARE HUGO PRIPP

E-post: [apripp@ous-hf.no](mailto:apripp@ous-hf.no)

Are Hugo Pripp er forsker og biostatistiker ved Oslo senter for biostatistikk og epidemiologi, Forskningsstøtteavdelingen, Oslo universitetssykehus. Han er professor II ved Fakultet for helsevitenskap, OsloMet – storbyuniversitetet.

Forfatteren har fylt ut ICMJE-skjemaet og oppgir ingen interessekonflikter.

---

Å matche eksponerte og ueksponerte individer på kjønn, alder eller andre observerte bakgrunnsvariabler kan instinktivt virke lurt, men også instinktive innfall bør gjennomtenkes.

En kohortstudie undersøker sammenhengen mellom ulike eksponeringer og senere utfall. For å studere sammenhenger i observasjonelle studier må man ta hensyn til eventuelle konfunderende variabler som påvirker assosiasjonen mellom eksponeringen og utfallet. En strategi for dette er å gjøre de eksponerte og ueksponerte gruppene like hverandre med hensyn til konfunderende variabler.

## Matching kan justere for konfundering

I kohortstudier – i motsetning til kasus-kontroll-studier – kan matchingen som sådan kontrollere for den konfunderende effekten av matchingsvariablene. Den kan fjerne sammenhengen mellom eksponeringen og faktoren(e) som er brukt til matchingen (1). I tabell 1 presenteres resultatet fra en hypotetisk studiepopulasjon. Alder er en konfunder. Den relative risikoen for alle i studiepopulasjonen er 1,5, men den relative risikoen til unge eller eldre er 3,0. I den matchete kohortstudien er 10 % av de eksponerte individene matchet med tilsvarende antall ueksponerte individer innad i hver aldersgruppe. Estimater fra alle data er dermed kontrollert for den konfunderende effekten av alder. Dette estimatet kan vi tolke som effekten av eksponeringen i det eksponerte utvalget, gitt at det ikke er andre konfunderende variabler eller kilder til systematiske feil. Med «ekte» data blir det sjelden så enkelt.

### Tabell 1

En matchet kohortstudie fra en hypotetisk studiepopulasjon. RR = relativ risiko,  $RR_{\text{alder}}$  = relativ risiko justert for alder med Poisson-regresjon eller stratifisert Mantel-Haenszel-analyse.

Studiepopulasjonen	Matchet kohortstudie
--------------------	----------------------

Studiepopulasjonen				Matchet kohortstudie				
	Kasuser	Ikke-kasuser	Totalt	Relativ risiko	Kasuser	Ikke-kasuser	Totalt	Relativ risiko
Unge								
Eksponert	600	39 400	40 000	RR = 3	60	3 940	4 000	RR = 3
Ueksponert	100	19 900	20 000		20	3 980	4 000	
Eldre								
Eksponert	400	7 600	8 000	RR = 3	40	760	800	RR = 3
Ueksponert	1 200	70 800	72 000		13	787	800	
Alle								
Eksponert	1 000	47 000	48 000	RR = 1,5	100	4 700	4 800	RR = 3
Ueksponert	1 300	90 700	92 000	RR <sub>alder</sub> = 3	33	4 767	4 800	

## En sannhet med forbehold

At matching alene er en effektiv og tilstrekkelig metode for å ta hensyn til konfundering i en virkelig kohortstudie, er lite sannsynlig. Da kan ingen andre konfunderende variabler eller systematiske feilkilder være til stede. Selv om man matcher på flere variabler, vil det ofte være andre observerte og uobserverte konfundere. I løpet av kohortens observasjonstid kan det bli forskjellig frafall og antall manglende observasjoner blant eksponerte og ueksponerte, noe som kan gi systematiske feil. I tillegg kan matching i kohortstudier være tidkrevende, særlig hvis man skal «koble sammen» eksponerte og ueksponerte individer for mange faktorer (2).

En viktig egenskap med utvalg i observasjonelle studier er å beskrive en studiepopulasjon (3). Ofte er det da en fordel å analysere et tilfeldig utvalg fra studiepopulasjonen, men det oppnår man ikke ved matching. Er eksponerte og ueksponerte individer i utvalget matchet for kjønn, kan vi ikke få et valid estimat av kjønnsfordelingen mellom disse i studiepopulasjonen.

## Statistisk analyse av en matchet kohort

Resultatene i tabell 1 synes å antyde at det ikke er nødvendig å utføre avanserte statistiske analyser i matchete kohortstudier, men dette er sjelden tilfellet i praksis. Hvis man ikke tar hensyn til variablene som er brukt til matching, viser metodisk forskning at resultatet kan bli systematisk feil (4). En anbefalt hovedregel er derfor å inkludere de matchete variablene i den statistiske analysen, f.eks. en multivariabel regresjonsmodell.

## Kan matching lønne seg?

Matching i kohortstudier kan i teorien gi en økt statistisk styrke slik at vi trenger færre individer i analysen. Det kan også virke «overbevisende» at eksponerte og ueksponerte er «like» med hensyn til bakgrunnsvariabler, men dette gir ingen garanti for en perfekt justering for konfundering eller andre systematiske feilkilder. Mer avanserte metoder for matching som tilbøyelighetsanalyser (*propensity score analysis*) kan være nyttige (5). Likevel bør man vurdere nøye hva man oppnår med matching sammenlignet med et tilfeldig utvalg der konfundering og andre systematiske feilkilder tas hensyn til i den statistiske analysen.

---

### LITTERATUR:

1. Kestenbaum B. Methods to control for confounding. *Epidemiology and biostatistics: an*

introduction to clinical research. New York, NY: Springer New York, 2009: 101–11.

2. Rothman K, Greenland S, Lash T. Matching. In: Rothman K, Greenland S, Lash T, red. Modern Epidemiology. 3. utg. Philadelphia, PA: Lippincott Williams & Wilkoms, 2008: 171–82.

3. Pripp AH. Populasjon og utvalg i statistikk. Tidsskr Nor Legeforen 2017; 137. doi: 10.4045/tidsskr.17.0124. [PubMed][CrossRef]

4. Sjölander A, Greenland S. Ignoring the matching variables in cohort studies - when is it valid and why? Stat Med 2013; 32: 4696–708. [PubMed][CrossRef]

5. Evans S. Commentary: Matched cohorts can be useful. BMJ 2003; 326: 357. [PubMed]

---

Publisert: 27. januar 2020. Tidsskr Nor Legeforen. DOI: 10.4045/tidsskr.19.0308

© Tidsskrift for Den norske legeforening 2020. Lastet ned fra tidsskriftet.no