

Aineiston keruun suunnittelu ja toteutus

Tero Vahlberg

Tutkimuksen vaiheet

1. Tutkimusongelma syntyy
- 2. Tutkimuksen suunnittelua**
- 3. Havaintoaineiston hankinta**
- 4. Tietojen talletus, muokkaus ja tarkastaminen**
5. Aineiston alustava tarkastelu ja tilastolliset analyysit
6. Tulosten tulkinta ja johtopäätökset
7. Tulosten esittäminen julkaisuissa tai tutkimusraporteissa

Tutkimusaineiston keruun suunnittelu

- Tutkimusaineiston keräämisen lähtökohtana on tutkimusongelma tai tutkimusongelmat
- Tutkimusongelma täsmentyy yleensä vaiheittain, kun tutkija perehtyy tutkimusaiheesta aiemmin tehtyihin tutkimuksiin ja olemassa oleviin teorian tietoihin kirjallisuuskatsauksen avulla
- Tutkimusongelman pitää olla selkeästi rajattu ja täsmällinen eikä tavoitteita saa olla liikaa yhtä tutkimusta kohti
- Selkeä tutkimusongelma on tärkeä tutkimuksen seuraavien vaiheiden eli tutkimusasetelman valinnan, aineiston keruun, tilastollisen analysoinnin ja julkaisujen kirjoittamisen kannalta
- Esimerkkejä tutkimusongelmista:
 - Onko kaatumisten ehkäisyinterventiolla vaikutusta vammahoitoa vaativien kaatumisten riskiin?
 - Mitkä tekijät ovat yhteydessä työntekijöiden työmotivaatioon?
 - Onko kahden lääkeryhmän välillä eroa hoitotuloksissa?

- Tutkimusongelman määrittämisen jälkeen mietitään:
 - Mikä on tutkimuksen havaintoyksikkö (potilas, perhe, sairaanhoitopiiri, kunta)?
 - Tutkimukseen sopiva tutkimusasetelma (esim. kokeellinen vs ei-kokeellinen, poikittainen vs pitkittäinen, poikkileikkaus vs tapaus-verrokki vs kohortti)
 - Arvioidaan tarvittavaa aineiston kokoa otoskokolaskelmien avulla
 - Mitä tietoja potilaista tarvitaan, jotta saadaan vastaukset tutkimuskysymyksiin? Tilastollisin termein potilaista mitattavia suureita tai ominaisuuksia kutsutaan muuttujiksi.
 - Millaista aineiston keruumenetelmää käytetään?

- Aineisto voidaan kerätä esim. seuraavilla tavoilla tai niitä yhdistelemällä:
 - Kyselylomakkeet: Valmiit strukturoidut lomakkeet vai tehdäänkö lomakkeet itse? Paperilomakkeet vai sähköisesti esim. Webropolilla vai www-kyselynä? Paperiset kyselylomakkeet voidaan muotoilla valmiiksi sellaiseen muotoon, että tiedot voidaan lukea skannauslaitteen avulla suoraan tietokantaan.
 - Haastattelut (vastaanotolla vai puhelimitse?)
 - Erilaiset päiväkirjat (esim. oireet päivittäin kirjattuna kahden viikon ajalta)
 - Potilasarkistot
 - Valmiit tilastot ja rekisterit:
 - Tilastokeskus, kunnat, KELA, THL, muut tutkimuslaitokset, jne.
 - Potilasrekisterit
 - Tutkittaville tehtävät kliiniset ja muut mittaukset (yleensä kokeellisissa tutkimuksissa)

Ohjeita kyselylomakkeen suunnitteluun

- Tee kohtuullisen pituisia ja ulkoasultaan selkeitä lomakkeita
- Mieti etukäteen tarkasti, mitkä ovat tutkimuksen kannalta oleellisia kysymyksiä – vältä liian pitkiä kyselylomakkeita
- Muotoile kysymykset mahdollisimman yksikäsitteisesti
- Käytä toisensa poissulkevia vastausvaihtoehtoja, jos ei ole kyse monivastauskysymyksestä
- "En osaa sanoa" tai "en tiedä" -vaihtoehtojen käyttöön ei ole olemassa yksiselitteisiä ohjeita, joten käytä niitä harkiten
- Numeerisena saatavana oleva tieto (esim. ikä/työkokemus vuosissa) kysytään sellaisenaan ilman valmista luokitusta. Mahdollinen luokitus voidaan tehdä myöhemmin aineiston analysoinnin yhteydessä
- Jos käytät VAS-muuttujia, niin varmista, että janan pituus kaikkien kysymysten kohdalla on tasan 100 mm

- Muista laittaa yksiköt näkyviin kysymyksiin, joihin annetaan vastaus numeroina (esim. aikaa mittaavat muuttujat vuosina tai kuukausina)
- Päivämäärät kysytään yhtenäisessä muodossa, esim. pp/kk/vvvv
- Kyselylomakkeesta tulee käydä selvästi ilmi, jos kysymykseen sallitaan vain yksi vastausvaihtoehto
- Käytä avoimia kysymyksiä harkiten (sisällönanalyysi on asia erikseen)
- Käytä lomakkeessa numeroituja vastausvaihtoehtoja tekstimuotoisten sijaan
- Seurantatutkimuksissa yhtenäiset lomakkeet. Kysymyksiä ei saa muuttaa kesken seurantatutkimuksen, muuten tietoja ei voida käyttää
- Älä tee muutoksia standardoituihin lomakkeisiin
- Testaa kyselylomaketta ennen kuin alat kerätä varsinaista aineistoa

Aineiston valmistelu ja tallennus

- Aineiston valmistelu ja tallennus ovat tärkeitä työvaiheita, jotta aineisto saadaan oikeanlaiseen muotoon tilastollista analysointia varten.

Tallennuspohja

- Tallennuspohja voidaan tehdä esim. Exceliin, tietokantaohjelmaan (Access) tai suoraan tilasto-ohjelmaan (SAS, SPSS)
- Tallennuspohjaan laitetaan AINA muuttujat sarakkeisiin ja havainnot riveille
- Muuttujien nimien tulee olla mahdollisimman lyhyitä ja selkeitä, eikä niissä saa esiintyä ääkkösiä (ä,ö,å), erikoismerkkejä (esim. %&#\$) tai välilyöntejä. Myös numeroita muuttujien nimissä on hyvä välttää.
- Aineistosta on hyvä luoda heti aluksi selkeä kuvaus esimerkiksi suoraan Excel-tiedostoon toiselle laskentasivulle tai erilliseen Word-dokumenttiin. Kuvauksesta tulee käydä ilmi muuttujien kuvaus (mitä asiaa muuttuja mittaa tai kuvaa) ja sisältö (luokkien selitykset esim. 1=nainen, 2=mies, jatkuvien muuttujien skaala).

- Tallennuspohjaan (tai erilliseen Word-dokumenttiin) kannattaa myös luoda erillinen laskentasiivu kommenteille, mihin voi kerätä tallentaessa ilmenneet ongelmat, kuten monen vastauksen kysymykset (jos vain yksi vastausvaihtoehto on sallittu), ristiriitaiset vastaukset tai ei-mahdolliset vastaukset

Tallennuspohjan rakenne

- Tallennuspohjan muotoilu on erittäin tärkeä vaihe, jotta aineistosta saadaan laadukas ja taataan sen jatkokäsiteltävyys
- Aineiston sekaan ei laiteta mitään ylimääräistä, kuten omia kommentteja, kysymysmerkkejä, Excelissä laskettuja tunnuslukuja, tyhjiä rivejä tai sarakkeita tai kuvia
- Älä laske Excelissä uusia muuttujia (esim. summamuuttujia tai keskiarvoja). Uudet muuttujat saa parhaiten laskettua tilasto-ohjelmassa ja Excelissä olevat kaavat vaikeuttavat aineiston käsittelyä
- Älä käytä piilotettuja rivejä tai sarakkeita

Aineiston tallentaminen

- Pyri tallentamaan mahdollisimman paljon suoraan numeeriseen muotoon, älä esimerkiksi täytä sukupuoli-sarakkeeseen ”mies” ja ”nainen”, vaan käytä numeerista koodausta (esimerkiksi nainen=1, mies=2). Tallennettaessa Exceliin varmista, että numeerisen muuttujan arvot tallentuvat numeerisessa muodossa
- Jos tallennat päivämääriä, tallenna ne aina samalla tavalla (esim. pp.kk.vvvv)
- Käytä aina samaa desimaalierotinta (piste tai pilkku)
- Tallenna muuttujien arvot pelkillä numeroilla, älä sekoita kirjaimia joukkoon
- Älä käytä värikoodeja, boldauksia tms. aineiston ryhmittelyssä. Mikäli aineistoa ryhmitellään jollain tavalla, on aineistossa oltava selkeä numeerinen muuttuja joka kertoo, mihin ryhmään mikäkin havainto kuuluu
- Puuttuva tieto jätetään tallentaessa aina tyhjäksi. Jos kuitenkin mielenkiinnon kohteena on erotella puuttuvan arvon syyt myöhemmin analysoinnissa, käytettävät numerot puuttuville havainnoille eivät voi olla mitään todellisia arvoja – puuttuvat päivämäärät jätetään aina tyhjiksi.
- Jos aineistoa tallentaa useampi ihminen, tee erillinen tallennusohje, jotta tiedostosta saadaan yhtenevät

Tallentamisen jälkeen

- Sosiaaliturvatunnukset, nimet, osoitteet ynnä muut tunnistetiedot säilytetään erillään muusta aineistosta – tee aineistosta tunnistevain, josta löytyy kutakin sosiaaliturvatunnusta vastaava juokseva ID-numerointi, jota käytetään myös lopullisessa tilasto-ohjelmaan siirrettävässä aineistossa. Lopullisessa aineistossa ei siis saa olla mukana ID-numeroinnin lisäksi muita tunnistetietoja.
- **SOSIAALITURVATUNNUKSIA, NIMIÄ TAI OSOITTEITA SISÄLTÄVÄÄ AINEISTOA EI SAA MISSÄÄN NIMESSÄ TOIMITTAA SÄHKÖPOSTITSE.**
- Muista säilyttää esim. syntymäaika iän laskemista varten
- Poista tilasto-ohjelmaan siirrettävästä aineistosta kaikki sellaiset muuttujat joita ei tulla käyttämään aineiston kuvailussa tai analyyseissä ja poista mahdolliset teksti-muotoiset muuttujat (esim. kyselylomakkeen kommentit)
- On tärkeää säilyttää myös paperiset lomakkeet vielä tallennuksen jälkeenkin
- Muista aineiston **VARMUUSKOPIOINTI**