

📄 Tutkimus- ja rekisteriaineistojen toisiokäyttö hoitotieteellisessä tutkimuksessa

KATJA JORONEN

TtT, professori
Hoitotieteen laitos
Turun yliopisto

OUTI KANSTE

TtT, dosentti, yliopistonlehtori
Lääketieteen tekniikan ja
terveystieteiden tutkimusyksikkö
Oulun yliopisto

KAISA MISHINA

TtT, erikoistutkija
Kliininen laitos,
Lastenpsykiatrian tutkimuskeskus
INVEST Lippulaiva
Turun yliopisto

TIIVISTELMÄ

Avoin tiede koskee yhä enenevässä määrin myös aineistojen avointa saatavuutta. Globaalisti ja kansallisesti kerätään yhä enemmän laadukkaita tutkimus- ja rekisteriaineistoja, joiden avulla voidaan vastata hoitotieteellisiin kysymyksenasetteluihin. Toisio- eli sekundaariaineistojen käyttö onkin lisääntynyt hoitotieteellisessä tutkimuksessa, mutta tutkimusaineistojen avaaminen jatkokäyttöön ja sekundaarianalyysit ovat edelleen melko harvinaisia. Tässä artikkelissa kuvataan tutkimus- ja rekisteriaineistojen toisiokäyttöä, tutkimusprosessia sekä arvioidaan toisiokäytön hyötyjä, rajoituksia ja eettisiä kysymyksiä terveys- ja hoitotieteellisessä tutkimuksessa. Lisäksi tässä kuvataan esimerkkejä tutkimus- ja rekisteriaineistojen toisiokäytöstä terveystieteellisissä tutkimuksissa.

Toisioaineistojen käyttö on taloudellista ja aikaa säästävää. Laajat väestökyselyt ja rekisterit mahdollistavat erilaisia otoskokoja ja otanta-asetelmiä. Toisaalta esimerkiksi rekistereistä saatava tieto mahdollistaa myös harvinaisten ilmiöiden tutkimisen. Toisioaineistot saattavat myös osoittaa uuden alkuperäistutkimuksen tarpeen. Toisioaineistojen rajoituksena on, että aineistojä arkistoidaan usein viiveellä eli ne voivat olla vanhoja. Aineisto saattaa myös sisältää puuttuvia tietoja tai tutkimuksen toteutus ja metodit ovat puutteellisesti kuvattuja. Toisioaineiston käyttö edellyttääkin tutkijalta perinpohjaista tu-

ABSTRACT

Secondary use of research and registry data in nursing science research

Katja Joronen, PhD, Professor
Outi Kanste, PhD, Docent, University lecturer
Kaisa Mishina, PhD, Senior researcher

Open science includes the open access to research data. Globally and nationally, there is an increasing number of high-quality research and registry data sets to answer questions in the field of nursing science. The use of secondary data has increased in nursing research, but the opening of research data for further use and secondary analysis remains quite rare. This article describes the secondary use of research and registry data, the research process, the evaluation of its benefits, limitations and ethical issues of research and registry data in health and nursing research. The article also describes examples of secondary use of research data and registry data in health studies.

The use of secondary materials is economical and time saving. Large population-based surveys and registers allow different sample sizes and sampling arrangements. On the other hand, information from registers also enables research on rare phenomena. The secondary analysis may also indicate the need for a new original study. The limitation of secondary data is that research materials are often archived with a delay, meaning they may be old. The data may contain missing information or the implementation and methods of the study are inadequately described. The use of secondary data requires the

Saapunut 31.01.2022

Hyväksytty julkaistavaksi 13.12.2022

tustumista aineistoon, sen laadun arviointia sekä eettisten kysymysten huomioimista.

Toisioaineistojen löydettävyyttä, saavutettavuutta sekä aineistojen uudelleen käyttöä terveystieteissä tulisi vahvistaa. Tämä edellyttää toisioaineistoihin ja sekundaarianalyysiin liittyvän menetelmäosaamisen vahvistamista jo perustutkintotasolla.

Avainsanat: avoin tiede, toisiokäyttö, tutkimusaineisto, rekisterit, väestötutkimus

researcher to have a thorough examination with the material, ethical issues, as well as quality assessment.

The findability, accessibility and re-use of secondary materials should be strengthened in nursing science. This requires the strengthening of methodological competence related to secondary analysis already at the undergraduate level of health sciences.

Key words: open science, dataset, registries, secondary data, population surveillance

Mitä tutkimusaiheesta jo tiedetään?

- Tutkimus- ja rekisteriaineistojen toisiokäyttö on lisääntynyt hoitotieteellisessä tutkimuksessa.
- Avoin julkaiseminen koskee yhä enenevässä määrin myös aineistojen avointa saatavuutta, ja monet tutkimuksen rahoittajat edellyttävät rahoittamansa hankkeen tuottamien tutkimusaineistojen avaamista.
- Tutkimus- ja rekisteriaineistojen toisiokäyttö voi lyhentää tutkimusprosessia merkittävästi.

Mitä uutta tietoa artikkeli tuo?

- Suomalaisessa palvelujärjestelmässä ja yhteiskunnassa kerätään systemaattisesti luotettavia aineistoja, joita voitaisiin hyödyntää hoitotieteellisessä tutkimuksessa.
- Artikkelit kuvaa kootusti kriteerejä tutkimus- ja rekisteriaineistojen toisiokäytön hyötyjen ja rajoitusten arviointiin.
- Artikkelissa tuodaan esille tutkimus- ja rekisteriaineistojen toisiokäytön tutkimusprosessia suhteessa alkuperäisaineistosta tehtävään tutkimukseen.

Mikä merkitys tutkimuksella on hoitotyölle, hoitotyön koulutukselle ja johtamiselle?

- Tutkimus- ja rekisteriaineistojen toisiokäyttöä voitaisiin hyödyntää vielä enemmän hoitotieteellisessä tutkimuksessa, hoitotyön kehittämis- ja innovaatiotoiminnassa, opetuksessa ja tiedolla johtamisessa.
- Tämän tutkimuksen menetelmällinen tieto voi rohkaista tutkijoita, terveysalan ammattilaisia, opettajia ja johtajia tutkimus- ja rekisteriaineistojen toisiokäytön aikaisempaa laajempaan hyödyntämiseen.

Johdanto

Toisioaineistoja hyödynnetään runsaasti monilla tieteenaloilla ja sähköiset järjestelmät sekä digitaaliset aineistot ovat tarjonneet uusia mahdollisuuksia toisioaineistojen hyödyntämiseen (O'Connor 2020). Hoitotieteessä toisioaineistojen käyttö on aiemmin ollut harvinaista (Dunn ym. 2015), mutta käyttö on kasvanut, erityisesti kansainvälisissä hoitotieteellisissä tutkimuksissa (O'Connor 2020). Aiemman toisioaineistojen vähäisen käytön on oletettu johtuvan muun muassa

siitä, että hoitotieteen tutkijat eivät olleet tietoisia tällaisten aineistojen saatavuudesta ja pääsystä niihin. Hoitotieteen tutkijoilla ei ole myöskään välttämättä riittävästi toisioaineistojen analyysitaitoja. Hoitotieteellinen kirjallisuus on huomionut yllättävän vähän toisioaineistoja osana tieteenalan metodologisia pohdintoja. (Dunn ym. 2015.) Kansainvälisen kyselytutkimuksen Suomen aineistoon perustuvien tulosten mukaan tieteenalakohteisessa vertailussa lääke- ja terveystieteiden edustajat näkivät eniten ongelmia aineistojen avoimessa jakamisessa. Erot tieteenalo-

jen välillä kertonevat erilaisten tutkimusmenetelmien ja -aineistojen lisäksi myös eri alojen kulttuureista ja perinteisistä tavoista suhtautua aineistoihin. (Enwald 2018.)

Toisioaineistot voidaan jakaa avoimen tieteen periaatteeseen pohjautuvaan **tutkimusaineistojen toisiokäyttöön** sekä **rekisteriaineistojen toisiokäyttöön** uusissa tutkimuksissa. *Avoin tiede* korostaa *tutkimusaineistojen toisiokäyttöä*. Avoin tiede voidaan jakaa kolmeen teemaan: FAIR-tutkimusdataan, avoimeen julkaisemiseen sekä avoimuuden kulttuuriin. FAIR-tutkimusdata tarkoittaa, että aineisto on löydettävä (*findable*), saavutettava (*accessible*), yhteentoimiva (*interoperable*) ja uudelleenkäytettävä (*reusable*). Näitä avoimuuden osa-alueita edellytetään sekä aineiston että metadatan eli aineiston kuvailutietojen osalta. Suomi on sitoutunut avoimen tieteen periaatteisiin ja Korkeakoulu- ja tutkimusyhteisön kansallisen linjauksen ja toimenpideohjelman (2021–2025) tavoitteisiin. Ohjelman osalinjaus, joka koskee tutkimusdatan avointa saatavuutta, on asettanut päämääräksi, että aineistot ja menetelmät ovat niin avoimia kuin mahdollista ja niin suljettuja kuin välttämätöntä. (Unifi 2018, Avoimen tieteen koordinaatio 2021.) Useat tutkimuksen rahoittajat, kuten Suomen Akatemia ja Euroopan komissio ovat linjanneet, että aineistojen käyttö tulee oletusarvoisesti mahdollistaa tutkimuksen päätyttyä (Tietoarkisto 2022, Suomen Akatemia 2022).

Rekisteriaineistojen toisiokäyttö uusissa tutkimuksissa perustuu Lakiin sosiaali- ja terveystietojen toissijaisesta käytöstä (552/2019), joka tuli voimaan Suomessa 1.5.2019. Kyseisen lain tavoitteena on ”*mahdollistaa sosiaali- ja terveydenhuollon toiminnassa syntyneiden asiakas- ja rekisteritietojen tietoturvallinen käsittely toisiokäytössä, sekä turvata yksilön luottamuksensuoja sekä oikeudet ja vapaudet henkilötietoja käsitellessä.*” Lain mukaan henkilötietojen toissijaisella käyttötarkoituksella tarkoitetaan henkilötietojen käsittelyä muussa käyttötarkoituksessa kuin henkilötietojen ensisijaisessa käyttötarkoituk-

sessä eli sellaisessa käyttötarkoituksessa, jossa henkilötiedot on alun perin tallennettu. Tällaisia toissijaisia käyttötarkoituksia ovat muun muassa tieteellinen tutkimus, tilastointi ja opetus. Samalla laki sääti sosiaali- ja terveysalan tietolupaviranomaisen, Findatan, joka ”*tekee muiden rekisterinpitäjien aineistoja koskevia tietolupapäätöksiä, vastaa päätöstensä mukaisten tietojen kokoamisesta, yhdistelystä, esikäsittelystä ja luovuttamisesta toissijaiseen käyttöön. Findata saa koota laissa säädettyä käyttötarkoitusta varten tietopyynnön perusteella eri rekisterinpitäjien henkilötietoja ja niitä yhdistelemällä tuottaa anonymoimia tietoja pyytäjän käyttöön.*” Findata on itsenäinen viranomainen, joka toimii Terveyden ja hyvinvoinnin laitoksen yhteydessä. (Laki sosiaali- ja terveystietojen toissijaisesta käytöstä 552/2019, Findata 2022.)

Sekä tutkimusaineistojen että rekisteriaineistojen toisiokäytössä voidaan erotella kahdenlaisia tutkimusaineistoja: aineistot, jotka ovat sisältävät henkilötietoja sekä aineistot, joissa näitä henkilötietoja ei ole. Edellistä säätelee tarkasti sekä Euroopan Unionin yleinen tietosuoja-asetus (GDPR), että Suomessa tätä asetusta täsmentävä ja täydentävä Tietosuojalaki (1050/2018).

Tämän artikkelin tarkoituksena on kuvata tutkimus- ja rekisteriaineistojen toisiokäyttöä ja sen tutkimusprosessia sekä arvioida toisiokäytön hyötyjä, rajoituksia ja eettisiä kysymyksiä terveys- ja hoitotieteellisessä tutkimuksessa. Lisäksi tarkoituksena on kuvata esimerkkejä tutkimus- ja rekisteriaineistojen toisiokäytöstä terveystieteellisissä tutkimuksissa. Tavoitteena on edistää tutkimus- ja rekisteriaineistojen toisiokäyttöä sekä toisioaineistojen löydettävyyttä, saavutettavuutta ja uudelleen käyttöä terveystieteissä.

Tutkimus- ja rekisteriaineistojen toisiokäytön muodot ja lähteet

Toisio- eli sekundaariaineiston analyysi tarkoittaa analyysia, joka kohdistuu jo olemassa olevaan aineistoon, ja joka testaa uusia

Taulukko 1. Esimerkkejä toisioaineistojen lähteistä.

<p>Julkisten toimijoiden aineistoja Sairaanhoitopiirit ja hyvinvointialueet, kunnat ja kaupungit Kansalliset tutkimuslaitokset (kuten Terveyden ja hyvinvoinnin laitos, Työterveyslaitos) Tilastokeskus (mm. kuolemansytilasto) Digi- ja väestötietovirasto (ent. Väestörekisterikeskus) Kansaneläkelaitos Eläketurvakeskus Oikeusrekisterikeskus Valvira, Fimea Opetushallitus Työ- ja elinkeinoministeriö Korkeakoulut ja muut oppilaitokset Aluehallintovirasto Tietoarkistot (kuten Findata, Tietoarkisto, Kanta) Syntymäkohortit (kuten kansalliset syntymäkohortit) Biopankit</p>
<p>Yksityisten ja kolmannen sektorin toimijoiden aineistoja Yksityiset sosiaali-, terveys- ja hoivapalvelut Järjestöt, yhdistykset ja seuraut Kuluttajadata Tutkimusryhmien hankkeet ja aineistot (esim. haastattelut, kyselyt, havainnoinnit)</p>
<p>Muita aineistoja Sosiaalinen media Älypuhelimet (mm. m-health) Genomitieto (Genomikeskus) Esineiden internet (IoT)</p>
<p>Kansainvälisiä aineistoja (esim. OECD, WHO, YK, CIMES*, DwB**, Eurostat***)</p>

* *Centralising and Integrating Metadata from European Statistics*, <https://cimes.casd.eu/>

** *Data without Boundaries*, <http://www.dwbproject.org/>

*** *European Statistics* <https://ec.europa.eu/eurostat>

hypoteeseja tai vastaa uusiin tutkimuskysymyksiin (Doolan & Froelicher 2009). Tästä analyysistä käytetään nimitystä sekundaari-analyysi. Toisioaineiston voidaan määrittellä olevan sellaista, joka on kerätty johonkin toiseen tarkoitukseen tai vastaamaan johonkin toiseen kysymykseen, kuin johon sitä sekundaarianalyysissä käytetään. Sekundaarianalyysin tekijä voi olla yhtäältä alkuperäisen tutkimusaineiston kerääjä tai sen keruusta täysin ulkopuolinen henkilö. (Cheng & Phillips 2014.) Alkuperäisen aineiston keruusta vastannut ryhmä voi esimerkiksi yhdistää eri aineistoja tai esittää aineistolle uusia kysymyksiä. Toisaalta toisioaineisto voi olla myös esimerkiksi terveydenhuollosta kertyvää rekisteriaineistoa, jota otetaan tutkimuskäyttöön, esittäen sille uusia tutkimuskysymyksiä tai -hypoteeseja. Tutkimusta toisioaineistosta voidaan myös tehdä yhdistämällä tut-

kimus- ja rekisteriaineistoja. Tämä tarkoittaa esimerkiksi kyselyillä kerättävien tietojen yhdistämistä rekisteriaineistojen tietoihin.

Tutkimusaineistojen toisiokäytön muodot tutkimuskäytössä voidaan jakaa kumuloituvaan, pysyvään ja julkiseen aineistoon. *Kumuloituva aineisto* syntyy tutkimuksen ollessa käynnissä, ja sen hallinnointi on mielekkäintä paikallisesti. *Pysyvä aineisto* liittyy tiettyyn tutkimustulokseen ja sitä on tärkeää säilyttää tulosten läpinäkyvyyden ja toistettavuuden toteutumiseksi. Pysyvän aineiston säilytys on luonteeltaan lähellä arkistointia ja se soveltuu koordinoitavaksi kansallisella tasolla. *Julkisen aineisto* voi olla joko tutkimuksen raaka-ainetta tai yksi sen tuotteista pysyvän aineiston ohessa. Julkinen aineisto liittyy tutkimusaineiston hallintaan silloin, kun sitä syntyy tutkimuksessa. Julkiset aineistot voivat tulla myös tiedeyhteisön ulkopuo-

Taulukko 2. Toisio- ja alkuperäisaineistosta tehtävän tutkimuksen yhtäläisyydet ja eroavaisuudet (mukaillen Doolan & Froelicher 2009).

Toisioaineistosta tehtävä tutkimus	Alkuperäisaineistosta tehtävä tutkimus
1. Aiheeseen perehtyminen esimerkiksi kirjallisuuskatsauksen menetelmin	
2. Tiedonaukon osoittaminen ja siihen vastaamisen mahdollisuuksien määrittäminen	
3. Tarvittavien lupien hankkiminen toisioaineistoon ja siinä käytettyihin tutkimusmenetelmiin perehtyminen	3. Tutkimuskysymysten muodostaminen ja otoksen ja tutkimusmenetelmien alustava suunnittelu
4. Tutkimuskysymysten määrittäminen	4. Tutkimussuunnitelman laatiminen, sisältäen tutkittavien rekrytoinnin, aineistonkeruun toteuttamisen ja eettisyyden suunnittelun
5. Alkuperäiseen tutkimukseen perehtyminen tarkemmin, esimerkiksi otoksen ja tutkimusasetelman osalta	5. Eettisen toimikunnan puoltavan lausunnon ja tutkimuslupien hakeminen
6. Aineistonhallintasuunnitelman tekeminen ja analyysin suunnittelu huomioiden myös eettiset näkökohdat	6. Tutkittavien rekrytointi
7. Tarvittavien lupien hankkiminen toisioaineiston käyttöön (ellei tarvittavia lupia ole hankittu jo aikaisemmin)	7. Aineistonkeruu
8. Aineiston analysointi	
9. Tutkimustulosten raportointi	

lelta, jolloin koordinaatiotarpeet kohdistuvat enemmän eettiseen hallintaan kuin pelkästään säilytykseen. (Unifi 2018, 4.)

Tutkimus- ja rekisteriaineistoja toisiokäyttöön on saatavilla useista eri lähteistä. Jotkut aineistot edellyttävät hakuprosessin aineiston saamiseksi, kun taas osa on julkisesti tai yksityisesti saatavilla. Aineistojen lähteitä ovat niin kansallisten ja kansainvälisten julkisten tahojen keräämät aineistot, yksityisten lähteiden aineistot kuin myös esimerkiksi digitaalisten lähteiden keräämät tiedot, kuten esimerkiksi big data. Big dataa tarkoitetaan erilaisia massadatoja, jotka on alun perin kerätty johonkin muuhun kuin tutkimustarkoitukseen (Sund 2015). Big dataa on esimerkiksi kuluttajadata, jota voidaan hyödyntää terveystieteellisessä tutkimuksessa. Esimerkiksi LoCard-tutkimuksessa käytettiin S-ryhmän asiakasomistajien ostotietoaineistoa muun muassa elintapojen tarkastelussa (Nevalainen ym. 2018, Erkkola ym. 2019). Taulukossa 1 on esitelty erilaisia toisioaineistojen tyypillisiä lähteitä.

Tutkimus- ja rekisteriaineistojen toisiokäytön tutkimusprosessi

Toisioaineistosta tehtävän tutkimuksen lähtökohdista voidaan käyttää joko tutkimuskysymystä tai itse aineistoa. Tutkimuskysymyksen toimiessa lähtökohdista, asetetaan ensin tutkimuskysymykset tai -hypoteesit ja etsitään sen jälkeen niihin vastaamiseen soveltuva aineisto. Jos taas lähtökohdista on aineisto, perehtyy tutkija tarkasti tiettyyn aineistoon ja sen sisältöön. Tämän jälkeen tutkija voi päättää millaiseen kysymykseen aineiston avulla voidaan vastata ja näin ollen muodostaa kysymyksen tai hypoteesin. Käytännössä kuitenkin on tyypillistä, että toisioaineistosta tehtävä tutkimus sisältää näitä molempia lähtökohdista iteratiivisena prosessina. Tutkija voi esimerkiksi muodostaa ensin laajan kysymyksen ja tarkentaa sitä aineiston avulla. (Cheng & Phillips 2014.)

Toisioaineistosta tehtävä tutkimus eroaa joiltain osin alkuperäisaineistosta tehdystä tutkimuksesta. Taulukossa 2 kuvataan vertai-

lu toisio- ja alkuperäisaineiston prosessista (mukaellen Doolan & Froelicher 2009). Toisioaineistoa hyödynnettäessä on tärkeää perehtyä siinä käytettyihin menetelmiin, huomioiden erityisesti tutkimuksen ja aineiston laatu ja luotettavuus. Mikäli laadussa on puutteita, tulee aineiston käyttöä harkita tarkoin.

Tutkimus- ja rekisteriaineistojen toisiokäytön hyödyt

Toisioaineistojen käyttö on monella tapaa hyödyllistä. Se on taloudellista ja aikaa säästävää, koska aineistonkeruuvaihe jää pois tai se lyhenee merkittävästi. (Dunn ym. 2015.) Aineiston arkistointia voidaan pitää tieteellisenä ansiona tutkijalle. Tämä siksi, että tutkimusjulkaisut, joiden aineisto on arkistoitu, saavat enemmän viittauksia kuin julkaisut, joiden aineisto ei ole saatavilla jatkokäyttöön. Tämä taas johtuu siitä, että kun arkistoituja aineistoja hyödynnetään uusiin tutkimuksiin, niissä viitataan aineiston lisäksi usein myös alkuperäisten tekijöiden aineistosta kirjoittamiin julkaisuihin. (Tietoarkisto 2022.) Toisioaineistot ovat usein laajoja, ja ne mahdollistavat pitkittäistutkimusasetelmia ja seuranta-tutkimuksia eli muutoksen tarkastelun eri aikapisteissä. Etuna on myös se, että esimerkiksi väestökyselyt ja rekisterit mahdollistavat erilaisia otoskokoja ja otanta-asetelmia, ja ne tarjoavat tutkijalle laajoja väestötason aineistoja. Toisaalta esimerkiksi rekistereistä saatava tieto mahdollistaa myös harvinaisten ilmiöiden tutkimisen.

Toisioaineistot ovat pääosin anonyymeja tai anonymisoituja, ja siten ne eivät aiheuta samalla tavalla riskiä tunnistettavuuteen tutkimukseen osallistuvalla (Doolan & Froelicher 2009). Etuna on myös se, että tutkijan ei tarvitse olla yhteydessä potilasiin tai asiakkaisiin, eikä tarvitse siis kontaktoida heitä.

Toisioaineistot ovat tyypillisesti niin sanotusti puhdistettuja aineistoja, jolloin niistä on pyritty saamaan pois esimerkiksi mahdolli-

set virheet, ja aineisto on mahdollisimman valmista käytettäväksi ja analysoitavaksi (Cheng & Phillips 2014). Myös tämä säästää osaltaan tutkijan aikaa. Lisäksi toisioaineistot mahdollistavat usein sen, että tutkimus on tarvittaessa helppo ja nopea toistaa.

Kysely- ja rekisteriaineistot mahdollistavat tietojen yhdistämisen laajasti uusilla ja innovatiivisilla tavoilla. Muuttujia ja eri osallistujaryhmiä voidaan tarkastella uudella tavalla, mikä rikastuttaa aineiston käyttöä. Tätä kautta toisioaineistot saattavat myös osoittaa uuden alkuperäistutkimuksen tarpeen ja aineiston keruulle tarkemman rajauksen. Esimerkiksi rekisteripohjaisissa syntymäkohorttitutkimuksissa (Kansalliset syntymäkohortit 1987 ja 1997) on yhdistetty lähes kaikki mahdolliset sosiaali- ja terveysrekisteri- sekä väestörekisteritiedot (THL 2021). Tällainen aineistojen yhdistäminen edellyttää, että alkuperäisessä aineiston keruussa on henkilötunnus käytössä. Kysely- ja rekisteriaineistojen yhdistäminen on omaleimaista Suomessa ja Pohjoismaissa, sillä vain harvassa muussa maassa toteutetaan näin järjestäytyntä ja kattavaa rekisteröintiä ja säännöllistä tiedonkeruuta. Toisioaineistojen käyttö luo ja vahvistaa tieteenalan sisäistä ja tieteenalojen välistä yhteistoimintaa, kuten esimerkiksi yhteistyötä muiden terveystieteiden, tilastotieteen, sosiologian ja psykologian tutkijoiden kanssa. Toisioaineistojen käytöllä voidaan myös välttää päällekkäiset tiedonkeruut. (Tietoarkisto 2022.)

Tutkimus- ja rekisteriaineistojen toisiokäytön rajoitukset

Toisioaineistoilla ja niiden käytöllä on myös rajoituksia. Aineistot kertyvät ja niitä arkistoidaan viiveellä eli ne voivat olla vanhoja. Toisaalta nämä aineistot voivat tarjota tärkeän historiallisen näkökulman tutkimusaiheeseen. Toisioaineistot saattavat myös mahdollistaa aiheen tarkastelun pitkällä aikajänteellä, jolloin on mahdollista

tarkastella ajallisia muutoksia. (Doolan & Froelicher 2009.)

Toisioaineistoissa saattaa olla runsaasti puuttuvia tietoja, esimerkiksi seurantatutkimuksissa, tai tutkimusmenetelmät, kuten esimerkiksi otanta, voivat olla heikkolaa-tuisia tai puutteellisesti kuvattuja (Doolan & Froelicher 2009). Toisaalta esimerkiksi rekisteriaineistoilla Suomessa on usein pa-rempi edustavuus kuin kyselytutkimuksilla. Toisioaineistoja käytettäessä tutkijan ei kuitenkaan ole välttämättä ollut mahdollista osallistua aineistonkeruuta koskeviin päätöksiin (Garmon 2007).

Toisioaineiston käyttö edellyttää tutkijalta perinpohjaista tutustumista aineistoon ja erityisesti siihen, mihin käyttötarkoitukseen aineisto on alun perin kerätty. Vasta tämän jälkeen tutkija voi tehdä päätöksen aineiston mahdollisesta käytöstä. Aineiston korkea laatu on sekundaarianalyysin edellytys ja sen mahdolliset rajoitukset tulisi kuvata selkeästi tutkimusraportissa. Lisäksi rajoituksena on se, että jo suunniteltu ja toteutettu tutkimus sekä aineistonkeruu ovat suunnattu alkuperäisten kysymysten vastaamiseen (Doolan & Froelicher 2009). Toisioaineistot lukitsevatkin tutkimuskysymykset tiettyihin uomiinsa; lisäkysymyksiä ei voida tehdä ja jokin tärkeä näkökulma saattaa puuttua aineistosta. Suurten sekundaariaineistojen käytössä joudutaan usein tekemään kompromisseja, jotta sekundaariaineiston tarjoamat mahdollisuudet ja tutkimuksen teoreettinen kysymyksenasettelu saadaan vastaamaan toisiaan eli edetään aineiston ehdoilla (Sund 2015).

Lisäksi esimerkiksi rekistereihin kertyy tietoa hyvin rajatusti, ja toisaalta esimerkiksi väestökyselyissä ei voida kysyä kaikkea, esimerkiksi hyvin arkaluonteisia aiheita. Tutkijan tulee tutustua tarkasti aineistoon ja sen metatietoihin sekä primaaritutkimuksen kysymyksiin, ennen kuin hän voi asettaa tutkimuskysymyksensä. Rekisteri- ja väestötutkimuksiin ja niistä saataviin laajoihin aineistoihin saattaa liittyä myös riski niin sanottuun

aineiston kalasteluun, jossa voi olla helppo löytää tilastollisesti merkitseviä tuloksia, joilla ei kuitenkaan ole suurta käytännön merkitystä. Toisioaineiston käytön haasteena, ainakin rekisteriaineistojen osalta, voi olla myös tiukentunut tietosuojalainsäädäntö ja käyttö-luvista aiheutuvien kustannusten nousu, jotka voivat hankaloittaa aineistojen saamista tutkimuskäyttöön (Lamminpää ym. 2020).

Tutkimus- ja rekisteriaineistojen toisiokäyttötutkimusten laadun arviointi

Toisioaineistojen laadun arviointi on ensiarvoisen tärkeää tulosten luotettavuuden varmistamiseksi. Ensiksikin on arvioitava aineiston relevanssia: vastaako aineisto kysymyksenasettelun vaatimuksiin. Tätä voidaan arvioida tutkimuksen käsitteiden, otoksen ja menetelmien kautta. Tutkimuksen käsitteiden tulee olla yhteneviä sekä alkuperäis- että sekundaaritutkimuksessa, ja niiden tulee vastata vallalla olevaa ymmärrystä ilmiöstä. Lisäksi otos tulee arvioida suhteessa uuteen kysymyksenasetteluun: onko otos edustava, käytetäänkö koko- vai osaotosta ja tullaanko aineistoa rajaamaan joltakin osin. (Doolan & Froelicher 2009.) Myös primaariaineiston menetelmiä tulee arvioida tarkasti: miten aineistoa on kerätty ja ovatko mittarit standardoituja ja validoituja - ja onko esimerkiksi mittaria päivitetty aineistonkeruun jälkeen. Laadullisessa tutkimuksessa on arvioitava esimerkiksi haastatteluteemojen relevanssi sekundaaritutkimuksen näkökulmasta.

Toisioaineistoja tulee arvioida myös ajan-kohtaisuuden kautta: milloin aineisto on kerätty ja miten aikaviive vaikuttaa kysymyksenasetteluun ja tuloksiin. Myös aineiston tarkkuuteen tulee kiinnittää huomiota. Tutkijan tulee tutustua tarkasti alkuperäiseen kyselylomakkeeseen, raakadataan, metadataan ja manuaaleihin. Tulee huomioida, onko aineisto eheä, puuttuuko sieltä vastauksia ja liittyykö aineistoon virheitä. Myös aineiston lähteen luotettavuus tulee arvioida. (Dunn ym. 2015.)

Eettiset kysymykset tutkimus- ja rekisteriaineistojen toisiokäyttötutkimuksissa

Toisioaineistoista toteutettaviin tutkimuksiin sovelletaan samoja eettisiä ohjeituksia kuin alkuperäisaineistoihin. Niiden tulee kunnioittaa tutkittavan ihmisarvoa, itsemääräämisoikeutta ja yksityisyyttä eikä tutkimuksesta saa aiheutua merkittäviä riskejä, vahinkoja tai haittoja. Luottamuksellisuus tulee säilyttää koko tutkimusprosessin ajan. (Brakewood & Poldrack 2013, TENK 2019.) Lisäksi Suomessa kaikessa tieteellisessä tutkimuksessa noudatetaan TENK:n (2019) laatimia ohjeita hyvästä tieteellisestä käytännöstä ja tietosuojalakia (1050/2018).

Mikäli kyseessä on yksilöihin kohdistuva tutkimus, tulee ensin selvittää, onko tutkimukselle ollut tarpeen ja onko tehty eettinen ennakoarviointi. Mikäli tutkimusasetelman eettisestä ennakoarvioinnista on säädetty erikseen lääketieteellisestä tutkimuksesta annetussa laissa (488/1999), on aineiston mahdollisesta toisiokäytöstä myös mainittu lausunnossa. Mikäli tutkimus ei kuulu ko. lain piiriin, mutta kohdistuu ihmiseen, on selvítettävä, onko tutkimukselle tehty organisaation eettinen ennakoarviointi tai ihmistieteiden eettinen ennakoarviointi. Tällaisia tutkimuksia on muun muassa yhteiskuntatieteiden ohella esimerkiksi ihmiseen kohdistuva ei-kajoava terveystai lääketieteellinen tutkimus. (TENK 2019.) Myös näissä tapauksissa lausunnossa on yleensä maininta aineiston mahdollisesta toisiokäytöstä. On huomioitava, että tutkimukselle on tehtävä eettinen ennakoarviointi, jos siihen sisältyy jokin seuraavista asetelmista: 1) tutkimukseen osallistumisesta poiketaan tietoon perustuvan suostumuksen periaatteesta, 2) tutkimuksessa puututaan tutkittavien fyysiseen koskemattomuuteen, 3) tutkimus kohdistuu alle 15-vuotiaisiin ilman huoltajan erillistä suostumusta tai informointia, jonka perusteella huoltajalla

olisi mahdollisuus kieltää lasta osallistumasta tutkimukseen, 4) tutkimuksessa tutkittaville esitetään poikkeuksellisen voimakkaita ärsykyksiä, 5) tutkimuksessa on riski aiheuttaa tutkittaville tai heidän läheisilleen normaalin arkielämän rajat ylittävää henkistä haittaa tai 6) tutkimuksen toteuttaminen voi merkitä turvallisuusuhkaa tutkittaville tai heidän läheisilleen. (TENK 2019.)

Mikäli alkuperäistutkimukselle ei ole tehty eettistä ennakoarviointia, on TENK:n linjaus, ettei ennakoarviointilausuntoa anneta aineistonkeruun jälkeen (TENK 2019). On kuitenkin huomioitava, että Lain sosiaali- ja terveystietojen toissijaisesta käytöstä tavoitteena on nimenomaan ”*mahdollistaa sosiaali- ja terveydenhuollon toiminnassa sekä sosiaali- ja terveysalan ohjaus-, valvonta-, tutkimus- ja tilastotarkoituksessa tallennettujen henkilötietojen tehokas ja tietoturvallinen käsittely sekä niiden yhdistäminen Kansaneläkelaitoksen, Väestörekisterikeskuksen, Tilastokeskuksen ja Eläketurvakeskuksen henkilötietoihin*” (Toisio-laki 552/2019, 1 §). Eli edellä mainittuja aineistoja voidaan hyödyntää tieteellisessä tutkimuksessa, vaikka niitä ei olisi alun perin tallennettu mainitussa tarkoituksessa. Samaan aikaan lain tavoitteena on ”*turvata yksilön luottamuksensuoja sekä oikeudet ja vapaudet henkilötietoja käsiteltäessä*” (Toisio-laki 552/2019, 1 §).

Tutkimuksesta saadun hyödyn näkökulmasta voidaan myös argumentoida, että aineiston toisiokäyttö voi tuottaa uusia tärkeitä tuloksia ja sitä kautta hyötyjä suuremmalle ihmisjoukolle. Mikäli jo kerätyn aineiston kautta voidaan vastata uusiin tutkimuskysymyksiin, on eettisesti arveluttavaa rasittaa tutkittavia keräämällä uusi aineisto tätä kysymyksenasettelua varten. Näin ollen sekundaarianalyysi voi myös vahvistaa tutkimuksen eettisyyttä. Voidaan jopa kysyä, voidaanko aineiston sekundaarianalyysien tekemistä jättämistä perustella eettisesti. (Brakewood & Poldrack 2013.)

Esimerkkejä tutkimus- ja rekisteriaineistojen toisiokäytöstä

Esimerkkejä tutkimusaineistojen toisiokäytöstä

Laajojen väestötutkimusten aineistojen hyödyntäminen hoitotieteellisessä tutkimuksessa on yhä yleisempää. Suomessa esimerkiksi Terveyden ja hyvinvoinnin laitos tuottaa laajoja väestötutkimuksia, jotka tuottavat tietoa koko väestön terveydestä sekä terveyteen ja hyvinvointiin vaikuttavista tekijöistä. Tällaisia ovat esimerkiksi kansallinen terveys-, hyvinvointi- ja palvelututkimus (FinSote), kansallinen FinTerveys-tutkimus, ulkomailla syntyneiden hyvinvointitutkimus (FinMonik), korkeakouluopiskelijoiden terveys- ja hyvinvointitutkimus (KOTT) sekä lasten ja nuorten terveysseurantatutkimus (LATE) (THL 2022a).

Esimerkkinä tässä artikkelissa käsittelemme Kouluterveyskyselyn aineiston käyttöä. Terveyden ja hyvinvoinnin laitoksen toteuttama Kouluterveyskysely tuottaa maakunnallista ja paikallista seurantatietoa erikäisten lasten ja nuorten hyvinvoinnista, terveydestä, koulunkäynnistä ja opiskelusta, osallisuudesta sekä avun saamisesta ja palvelujen tarpeisiin vastaavuudesta. Kyselyyn osallistuvat perusopetuksen 4. ja 5. vuosiluokan oppilaat, perusopetuksen 8. ja 9. vuosiluokan oppilaat, lukiokoulutuksen 1. ja 2. vuoden opiskelijat, ammatillisten oppilaitosten 1. ja 2. vuoden perustutkintoa opiskelevat alle 21-vuotiaat opiskelijat. Vastaajia on yhteensä lähes 300 000, eikä aineisto sisällä henkilötunnusta. Kouluterveyskysely toteutetaan joka toinen vuosi. Tietoja on kerätty perusopetuksen 8. ja 9. luokan oppilailta vuodesta 1996 ja lukioissa vuodesta 1999. Ammatillisissa oppilaitoksissa opiskelevat ovat mukana vuodesta 2008 alkaen, ja perusopetuksen 4. ja 5. luokkaa käyvät lapset vuodesta 2017 alkaen. Perusopetuksen 4. ja 5. vuosiluokan oppilaiden huoltajat ovat vastanneet kyselyyn vuosina 2017 ja 2019. Tutkimukseen osallistuminen on vapaaeh-

toista, ja oppilaat vastaavat joko sähköisiin tai paperisiin lomakkeisiin luokkahuoneessa opettajan ohjauksessa. (THL 2022b.)

Kouluterveyskyselyaineiston avulla on vastattu lukuisiin hoitotieteellisiin kysymyksenasetteluihin, kuten esimerkiksi koululaisten terveyden ja oireilun määrittäjiin (Wargh ym. 2015, Karjalainen ym. 2017, Jaakkola ym. 2019, Talja ym. 2019). Aineistojen etuna on se, että väestötutkimus on kansallisesti edustava ja se mahdollistaa monipuolisia monimuuttuja-analyyseja, ja siten usean tekijän yhteyksien tarkastelun samanaikaisesti ja eri muuttujien vaikutusten vakioinnin. Esimerkiksi perheeseen liittyvät tekijät, kuten keskusteluvaikeudet vanhemman kanssa, todettiin olevan yhteydessä yläkoululaisten masennusoireiluun (Wargh ym. 2015) ja ammattiin opiskelevien ahdistusoireiluun (Karjalainen ym. 2017), vaikka iän, sukupuolen ja perherakenteen vaikutus oli vakioitu. Taljan ym. (2019) tutkimustulokset osoittivat, että alakoulun 4.- ja 5.-luokkalaisten elämänmuutokset ja niiden kumuloituminen olivat yhteydessä runsaampaan masennusoireiluun muiden taustamuuttujien vakioinnin jälkeinkin. Jaakkolan ym. (2019) tuloksissa korostui, että ammattiin opiskelevien koettu terveydentila, oireilu ja terveyskäyttäytyminen oli hyvin erilaista eri koulutusaloilla.

Kouluterveyskyselyn aineiston kautta on tutkittu myös nuorten terveyskäyttäytymistä (Mure ym. 2014, Vellonen ym. 2015, Mason ym. 2017, Aho ym. 2018). Esimerkiksi Mureen ym. (2014) tulokset osoittivat, että perheen yhteisen aterian syöminen oli yhteydessä yläkoululaisten vähäisempään tupakointiin tai nuuskaamiseen, alkoholinkäyttöön ja laittomien huumeiden kokeiluun. Ahon ym. (2018) tutkimus osoitti, että perheen osallisuus nuoren elämään oli yhteydessä vähäisempään tupakointiin ammattiin opiskelevilla nuorilla. Kouluterveyskyselyn avulla on analysoitu myös kouluterveydenhuollon palvelujen käyttöä (Isosomppi ym. 2017, Kanste 2019, Välkkiä ym. 2020, Kosunen ym. 2021, Suoniemi ym. 2021). Välkki-

län ym. (2020) tutkimus osoitti, että erityisesti perheen heikommalla taloudellisella tilanteella sekä nuoren ulkomaalaistaustaisuudella oli yhteys koettuun tyydyttymättömään palvelutarpeeseen kouluterveydenhoitajan palveluissa sekä yläkouluissa, lukioissa että ammatillisissa oppilaitoksissa. Kansteen (2019) tutkimus toi esille, että tyydyttymättömään palvelutarpeeseen kouluterveydenhoitajan palveluissa oli yhteydessä nuoren toimintarajoitteisuus ja Suoniemen ym. (2021) tutkimuksen mukaan koulunkäynnin vaikeudet yläkoulussa.

Laadullisia aineistoja voidaan käyttää toisioaineistoina esimerkiksi yhdistämällä saman tutkimuksen eri aineistoja tai aineistoja eri tutkimuksista. Esimerkiksi Nurmi ym. (2017) yhdisti toisioaineistoon perustuvassa tutkimuksessaan kaksi aiemmin julkaistua laadullista hoitotyön johtajien haastatteluaineistoa ja esitti tämän jälkeen yhdistetylle aineistolle uusia tutkimuskysymyksiä. Aineistojen yhdistäminen mahdollisti aiheen laajemman ja syvemmän tarkastelun, korostaen raportoimattomia löydöksiä. Tobiano ym. (2021) puolestaan yhdisti julkaistut potilaiden ja hoitohenkilökunnan haastatteluaineistot ja suoritti niille sen jälkeen vertailevaa analyysia. Tällä tavoin he saivat vastattua kysymykseen, johon ainoastaan yhdellä alkuperäisaineistolla ei olisi ollut mahdollista vastata.

Interventiotutkimuksissa voidaan hyödyntää sekundaarianalyysia, esimerkiksi intervention implementoinnin osatekijöistä, jotka havaitaan intervention aikana tai sen jälkeen, tai tulosmuuttujien uudelleen tarkastelun kautta. Interventiotutkimuksissa kerätäänkin usein aineistoa, jota voidaan käyttää vastaamaan uusiin asetettuihin tutkimuskysymyksiin. Esimerkiksi Friedman ym. (2014) toteuttivat sekundaarianalyysin kotisairaanhoidajien käyntien intervention vaikuttavuudesta ikääntyneiden päivittäisiin toimintoihin (ADL) siten, että he tutkivat tulosmuuttujia yksitellen alkuperäisen tutkimuksen summamuuttujan keskiarvon sijas-

ta. Tutkimus osoitti tarkemmin, mitkä päivittäiset toiminnot vahvistuivat ja mitkä heikkenivät intervention seurauksena.

Esimerkkejä rekisteriaineistojen toisiokäytöstä

Rekisteriaineistot ovat tyypillisiä esimerkkejä toisioaineistosta. Rekisteriaineistojen käyttö hoitotieteellisessä tutkimuksessa mahdollistaa laajojen, koko väestötason aineistojen hyödyntämisen. Sosiaali- ja terveysrekisteriaineistot koostuvat usein erilaisista sähköisistä sairaaloiden ja tutkimuslaitosten ylläpitämistä terveys- ja potilasrekistereistä. (Lamminpää ym. 2020.) Tilastot ja rekisterit tarjoavatkin laajasti tietoja sosiaali- ja terveydenhuollon eri aihealueilta päätöksenteon, kehittämisen ja tutkimuksen tueksi (THL 2022c). Suomessa sosiaali- ja terveysalan tietolupaviranomainen Findata huolehtii käyttöluvista sosiaali- ja terveystietojen toissijaiseen käyttöön, mikäli niitä tarvitaan usealta eri sote-rekisterinpitäjältä tai yksityiseltä sektorilta. Käyttöluvan saamiseksi tutkijan on haettava Findatalta tietolupaa. (Findata 2022.)

Toisioaineistoina käytettävät rekisterit voivat sisältää henkilötietoja, jolloin tietosuojan turvaaminen on erityisen tärkeää. Henkilötiedolla tarkoitetaan sellaisia tietoja, joiden avulla henkilö on mahdollista tunnistaa joko suoraan tai välillisesti. Suoran tunnistamisen henkilötietoja on esimerkiksi nimi ja henkilötunnus ja välillisen puolesta esimerkiksi sukupuoli, ikä ja koulutus. Välillisiä tietoja yhdistämällä voi olla mahdollista tunnistaa henkilö. (Findata 2022.) On huomioitavaa, että tunnisteellisia tietoja luovutetaan tutkijoille ainoastaan silloin, kun käyttötarkoitus sitä välttämättä edellyttää. Tiedot luovutetaan ensisijaisesti anonymina tai pseudonymina ja tietojen saajaa koskee salassapitovelvollisuus. (Sosiaali- ja terveysministeriö 2022.) Henkilötietoja sisältävien aineistojen käsittely, säilytys sekä hävittäminen, anonymisointi tai arkistointi tulee tapahtua asianmukaisesti.

ti ja tietoturvallisesti (Tietosuojavaltuutetun toimisto 2022). GDPR sääntelee henkilötietojen käsittelyä, ollen keskeinen henkilötietoja sisältävissä tutkimuksissa. Tietosuojaperiaatteiden mukaan henkilötietoja on käsiteltävä lainmukaisesti, asianmukaisesti ja rekisteröidyn kannalta läpinäkyvästi. Aineistoa saa kerätä ja käsitellä vain nimenomaisesti tarkoitusta varten ja vain tarpeellisen määrän henkilötietojen käsittelyn tarkoitukseen nähden. Aineiston säilytyksen on tapahduttava siten, että henkilö on tunnistettavissa ainoastaan niin kauan kuin on tarpeen tietojenkäsittelyn tarkoitusten toteuttamista varten. Rekisterinpitäjän on tarvittaessa pystyttävä osoittamaan, että se noudattaa kaikkia vaadittuja tietosuojaperiaatteita. (Tietosuojavaltuutetun toimisto 2022.)

Suomessa Hoitoilmoitusjärjestelmä Hilmo tarjoaa tutkijoille mahdollisuuden väestön terveyden ja terveyspalveluiden käytön arviointiin toisioaineiston avulla. Hilmoon kerätään tietoa muun muassa väestön palvelujen käytöstä, hoitopäätöksistä, terveysongelmista ja epidemioista, terveyden edistämisen palveluista sekä ammattilaisten työnjaosta ja hoitokäytännöistä. Hilmo sisältää kolme eri osaa: sosiaalihuollon hoitoilmoitusrekisterin, (sosiaali-Hilmo), terveydenhuollon hoitoilmoitusrekisterin (terveys-Hilmo) sekä perusterveydenhuollon avohoidon, johon kuuluu myös työterveyshuolto ja kotihoito (Avohilmo). Hilmoon kerättyjä tietoja on hyödynnetty laajasti erilaisissa tutkimuksissa. Hoitotieteellisistä tutkimuksista esimerkiksi Välimäki ym. (2019) arvioivat pakkotoimien määrää ja niiden käytön ajallisia muutoksia psykiatrisissa sairaaloissa Suomessa, käyttäen tiedonlähteenä Hilmon rekisteritietoa.

Toinen esimerkki rekisteriaineistosta ovat terveydenhuollossa käytettävät vaaratapahtumien ilmoitusjärjestelmät. Niiden tarkoituksena on edistää potilaiden ja henkilökunnan turvallisuutta rekisteröimällä ja sitä myötä lisäämällä ymmärrystä tapahtuneista vaaratapahtumista. Ne auttavat hoitohenkilökuntaa ennakoimaan ja ennaltaehkäise-

mään vaaratapahtumia, lisäämällä tietoa tapahtuneista vaaratapahtumista, mikä taas voi johtaa parempaan vaaratilanteiden hallintaan tulevaisuudessa (Anderson ym. 2013).

Yksi esimerkki tällaisesta työ- ja potilasturvallisuusilmoitusjärjestelmästä on HaiPro, joka on internetpohjainen raportointijärjestelmä ilmoitusten tekemiseksi ja sitä käytetään systemaattiseen turvallisuustapahtumien raportointiin terveydenhuollon organisaatioissa (Awanic 2021). Terveydenhuollossa kerättyä HaiPro-aineistoa on jonkin verran hyödynnetty myös tutkimuskäytössä toisioaineistona. Sen avulla on esimerkiksi tutkittu väkivaltaan liittyvien vaaratapahtumien esiintymistä oikeuspsykiatrisilla osastoilla (Kuivalainen ym. 2014, Kuosmanen ym. 2022), tiedonhallinnan vaaratapahtumia sairaaloissa (Jylhä ym. 2016a, 2016b) sekä lääkehoidon vaaratapahtumia (Härkänen ym. 2020) ja lääkityspoikkeamiin liittyvää kommunikaatiota (Syyrilä ym. 2020) sairaaloissa.

Pohdinta ja johtopäätökset

Avoin tiede edellyttää aineistojen avaamisen uudelleen tutkimuskäyttöön. Sähköiset aineistot mahdollistavat sen, että tutkimus- ja rekisteriaineistojen toisiokäyttöä voidaan hyödyntää yhä enemmän ja monipuolisemmin. Tutkimus- ja rekisteriaineistojen toisiokäytöllä on useita hyötyjä. Ne mahdollistavat laajojen väestötason aineistojen hyödyntämisen ja aineistojen innovatiivisen yhdistämisen. Lisäksi niiden käyttö on taloudellista ja aikaa säästävää.

Tutkimus- ja rekisteriaineistojen toisiokäytöllä on kuitenkin myös rajoituksia. Aineistot eivät välttämättä ole ajantasaisia, ja niissä saattaa olla runsaasti puuttuvia tietoja. Ne ovat myös tyypillisesti kerätty johonkin muuhun tarkoitukseen kuin siihen, johon niitä aineistojen toisiokäyttötutkimuksissa hyödynnetään. Tällöin tutkimusmenetelmälliset ja tutkimuksen luotettavuuteen liittyvät haasteet ovat mahdollisia.

Toisiokäytettävä tutkimus- ja rekisteriaineisto on lähtökohdiltaan joko sellaista, joka sisältää henkilötietoja tai sellaista, jossa koko tutkimusaineisto alusta alkaen sisältää täysin anonyymiä tietoa. Mikäli tutkimusaineistossa käsitellään henkilötietoja, se edellyttää tietosuojalainsäädännön huomioon ottamista tutkimusprosessin kaikissa vaiheissa. Rekisteriaineistoista säätää laki sosiaali- ja terveystietojen toissijaisesta käytöstä (552/2019).

Myös tutkijat itse voivat avata keräämiään aineistoja. Tällaiset aineistojen jatkokäyttömahdollisuudet olisi hyvä huomioida esimerkiksi aineistohallintasuunnitelmassa, tutkimuksesta viestimisessä, varmistamalla asianmukaiset aineistojen kuvailutiedot sekä hyödyntämällä avoimia alustoja. Tämän mahdollistamiseksi tutkijalla tulisi olla riittävästi tietoa ja taitoja.

Tutkimus- ja rekisteriaineistojen toisiokäytön mahdollisuuksia on runsaasti tarjot-

la myös hoito- ja terveystieteelliseen tutkimukseen, ja niiden käyttöön on tarpeen saada menetelmäkoulutusta ja osaamisen vahvistamista jo perustutkintotasolla. Aineiston toisiokäytössä tutkijan on tärkeää perehtyä huolellisesti käytettävään aineistoon ja siihen, miten se on kerätty ja mihin käyttötarkoituksiin. Lisäksi tutkijan tulee huomioida millaisia luotettavuutta heikentäviä rajoituksia aineistoon voi liittyä. Tutkijan on tärkeä suunnitella tutkimus huolellisesti ja tarvittaessa tehdä yhteistyötä primaariaineiston keränneiden tutkijoiden kanssa.

VASTUUALUEET

Tutkimuksen suunnittelu: KJ, OK, KM, aineistonkeruu: KJ, OK, KM, aineiston analysointi: KJ, OK, KM, käsikirjoituksen kirjoittaminen: KJ, OK, KM, käsikirjoituksen kommentointi: KJ, OK, KM.

LÄHTEET

- Aho H., Koivisto AM., Paavilainen E. & Joronen K. (2018) Parental involvement in their adolescents' lives and adolescent smoking. *Health Promotion International* **33**(5), 846–857.
- Anderson JE., Kodate N., Walters R. & Dodds A. (2013) Can incident reporting improve safety? Healthcare practitioners' views of the effectiveness of incident reporting. *International Journal for Quality in Health Care* **25**(2), 141–150.
- Avoimen tieteen koordinaatio, Tieteellisten seurain valtuuskunta. (2021) Tutkimusaineistojen ja -menetelmien avoimuus. Korkeakoulu- ja tutkimusyhteisön kansallinen linjaus ja toimenpideohjelma 2021–2025. Osalinjaus 1: Tutkimusdatan avoin saatavuus. Susanna Nykyri, Selma Päällysaho, Tomi Rosti (pj.), Anne Sunikka, Anssi Neuvonen, Mari Elisa Kuusniemi (toim.). Vastuullisen tieteen julkaisusarja 5:2021. Tiedonjulkistamisen neuvottelukunta ja Tieteellisten seurain valtuuskunta.
- Awanic Oy. (2021) Haipro. <https://awanic.fi/eng/products-and-services/#haipro> (14.10.2022)
- Brakewood B. & Poldrack RA. (2013) The ethics of secondary data analysis: Considering the application of Belmont principles to the sharing of neuroimaging data. *NeuroImage* **82**(15), 671–676.
- Cheng HG. & Phillips MR. (2014) Secondary analysis of existing data: opportunities and implementation. *Shanghai Archives of Psychiatry* **26**(6), 371–375.
- Doolan DM. & Froelicher ES. (2009) Using an existing data set to answer new research questions: A methodological review. *Research and Theory for Nursing Practice: An International Journal* **23**(3), 203–215.
- Dunn SL., Arslanian-Engoren C., DeKoekkoek T., Jadack R. & Scott LD. (2015) Secondary data analysis as an efficient and effective approach to nursing research. *Western Journal of Nursing Research* **37**(10), 1295–1307.
- Enwald H. (2018) Tutkimusaineiston avoin jakaminen – tutkimusorganisaatioiden jäsenten käsityksiä, kokemuksia ja mielipiteitä. *Informaatiotutkimus* **37**(4), 26–47.
- Erkkola M., Fogelholm M., Saarijärvi H., Uusitalo L. & Nevalainen J. (2019) Kuluttajadatan mahdollisuudet ja haasteet kansanterveystutkimuksessa; case LoCard. *Sosiaalilääketieteellinen aikakauslehti* **56**(2), 76–87.
- Findata. (2022) <https://findata.fi/> (26.10.2022)
- Friedman B., Li Y., Liebel DV. & Powers BA. (2014) Effects of a home visiting nurse intervention versus care as usual on individual activities of daily living: a secondary analysis of a randomized controlled trial. *BMC Geriatrics* **14**(24), 1–13.
- Garmon SC. (2007) Issues associated with secondary analysis of population health data. *Applied Nursing Research* **20**(2), 94–99.

- Härkänen M., Turunen H. & Vehviläinen-Julkunen K. (2020) Differences Between Methods of Detecting Medication Errors: A Secondary Analysis of Medication Administration Errors Using Incident Reports, the Global Trigger Tool Method, and Observations. *Journal of Patient Safety* **16**(2), 168–176.
- Isosomppi M., Rantanen A., Kivimäki H., Koivisto AM. & Joronen K. (2017) Ammattiin opiskelevien nuorten opiskeluohjelman palvelujen käyttö ja tupakointikäyttäytyminen. *Hoitotiede* **29**(2), 125–138.
- Jaakkola J., Rantanen A., Luopa P., Koivisto AM. & Joronen K. (2019) Self-rated health, symptoms and health behaviour of upper secondary vocational students by field of study. *Scandinavian Journal of Caring Sciences* **33**(1), 144–155.
- Jylhä V., Bates DW. & Saranto K. (2016a) Adverse events and near misses relating to information management in a hospital. *Health Information Management Journal* **45**(2), 55–63.
- Jylhä V., Bates DW. & Saranto, K. (2016b) Critical factors in the information management process: the analysis of hospital-based patient safety incident reports. *Finnish Journal of EHealth and EWellfare* **8**(4), 164–176.
- Kanste O. (2019) Subjective unmet needs for school health services among adolescents with different disabilities: a population-based study in Finland. *International Journal of Caring Sciences* **12**(3), 1346–1355.
- Karjalainen V., Rantanen A., Matikka A., Koivisto AM. & Joronen K. (2017) Perhetekijöiden ja sosioekonomisten tekijöiden yhteys ammattiin opiskelevien nuorten itsearvioituun ahdistuneisuuteen. *Hoitotiede* **29**(2), 113–124.
- Kosunen S., Rantanen A., Koivisto AM., Ikonen R. & Joronen K. (2021) Seksuaalista häirintää ja väkivaltaa kokeneiden nuorten taustatekijöiden yhteys tuen ja avun saamiseen koulun aikuisilta. *Nuorisotutkimus* **39**(4), 36–50.
- Kuivalainen S., Vehviläinen-Julkunen K., Putkonen A., Louheranta O. & Tiihonen J. (2014) Violent behaviour in a forensic psychiatric hospital in Finland: An analysis of violence incident reports. *Journal of Psychiatric and Mental Health Nursing* **21**(3), 214–218.
- Kuosmanen A., Tiihonen J., Repo-Tiihonen E. & Turunen H. (2022) Voluntary patient safety incidents reporting in forensic psychiatry-What do the reports tell us? *Journal of Psychiatric and Mental Health Nursing* **29**(1), 36–47.
- Laki lääketieteellisestä tutkimuksesta (488/1999). <https://www.finlex.fi/fi/laki/ajantasa/1999/19990488> (14.10.2022)
- Laki sosiaali- ja terveystietojen toissijaisesta käytöstä (552/2019). <https://www.finlex.fi/fi/laki/alkup/2019/20190552> (14.10.2022)
- Lamminpää R., Gissler M. & Vehviläinen-Julkunen K. (2020) Rekisteriaineistot ja niiden hyödyntäminen hoitotieteen tutkimuksessa - esimerkkinä raskausdiabetesta sairastavat 35 vuotta täyttäneet synnyttäjät. *Hoitotiede* **32**(20), 15–24.
- Mason A., Rantanen A., Kivimäki H., Koivisto AM. & Joronen K. (2017) Family factors and health behaviour of thin adolescent boys and girls. *Journal of Advanced Nursing* **73**(1), 177–189.
- Mure K., Konu A., Kivimäki H., Koivisto AM. & Joronen K. (2014) Perheaterioinnin yhteys 8.- ja 9.-luokkalaisten päihteidenkäyttöön. *Sosiaalilääketieteellinen aikakauslehti* **51**(2), 88–100.
- Nevalainen J., Erkkola M., Saarijärvi H., Näppilä T. & Fogelholm M. (2018). Large-scale loyalty card data in health research. *Digital Health* **4**, 1–10.
- Nurmi SM., Halkoaho A., Kangasniemi M. & Pietilä AM. (2017) Collaborative partnership and the social value of clinical research: a qualitative secondary analysis. *BMC Medical Ethics* **18**(57), 1–12.
- O'Connor S. (2020) Secondary data analysis in nursing research: a contemporary discussion. *Clinical Nursing Research* **29**(5), 279–284.
- Sosiaali- ja terveysministeriö. (2022) *Usein kysytyt sote-tietojen toissijaisesta käytöstä*. <https://stm.fi/useinkysytyt-sote-tiedon-toisiokaytosta> (26.10.2022)
- Sund R. (2015) Miksi isoon dataan hukutaan? *Tieto & trendit – talous- ja hyvinvointikatsaus* (2), 40–45.
- Suomen Akatemia. (2022) Aineistonhallintasuunnitelma. <https://www.aka.fi/tutkimusrahoitus/hae-rahoitusta/nain-haet-rahoitusta/ohjehakemisto/aineistonhallinta/aineistonhallintasuunnitelma/> (14.10.2022)
- Suoniemi S., Rantanen A., Koivisto AM. & Joronen K. (2021) Self-reported school difficulties and the use of the school nurse services by adolescent students. *Children* **28**(8), 647.
- Syyrilä T., Vehviläinen-Julkunen K. & Härkänen M. (2020) Communication issues contributing to medication incidents: Mixed-method analysis of hospitals' incident reports using indicator phrases based on literature. *Journal of Clinical Nursing* **29**(13-14), 2466–2481.
- Talja T., Rantanen A., Koivisto AM. & Joronen K. (2019) Itsearvioitu masennusoireilu ja elämäntapahtumat 4. ja 5. luokan oppilailla – Kouluterveyskysely. *Sosiaalilääketieteellinen aikakauslehti* **56**(3), 168–180.
- TENK. (2019) *Ihmiseen kohdistuvan tutkimuksen eettiset periaatteet ja ihmistieteiden eettinen ennakkoarviointi Suomessa*. Tutkimuseettisen neuvottelukunnan ohje 2019. PDF-julkaisu. <https://tenk.fi/fi/eettinen-ennakkoarviointi/ihmistieteiden-eettinen-ennakkoarviointi> (14.10.2022)
- THL. (2021) *Kansalliset syntymäkobortit 1987 ja 1997*. www.thl.fi/fbc1987 (14.10.2022)
- THL. (2022a) *Tulosraportit väestötutkimuksista ja rekisteriaineistoista*. <https://www.terveytemme.fi/raportit/index.html> (14.10.2022)
- THL. (2022b) *Tilastot ja data*. <https://thl.fi/fi/tilastot-ja-data> (14.10.2022)
- THL. (2022c) *Kouluterveyskysely*. <https://thl.fi/fi/tutkimus-ja-kehittaminen/tutkimukset-ja-hankkeet/kouluterveyskysely> (14.10.2022)
- Tietoarkisto. (2022) *Miksi aineistonhallintaa ja jatkokäyttöä?* <https://www.fsd.tuni.fi/fi/palvelut/aineistonhallinta/miksi-aineistonhallintaa-ja-jatkokaytto/> (14.10.2022)
- Tietosuojalaki (1050/2018). <https://www.finlex.fi/fi/laki/ajantasa/2018/20181050> (14.10.2022)
- Tietosuojavaltuutetun toimisto. (2022) *Tieteellinen tutkimus ja tietosuojat*. <https://tietosuoja.fi/tieteellinen-tutkimus> (26.10.2022)

- Tobiano G., Marshall AP. & Chaboyer W. (2021) Comparing Perceptions of Patient Nonparticipation in Nursing Care: A Secondary Analysis. *Journal of Nursing Scholarship* **53**(4),449–457.
- Unifi. (2018) *Avoin tiede ja data -toimenpideohjelma suomalaiselle tiedeyhteisölle*. Suomen yliopistojen rehtorineuvosto UNIFI ry. PDF-julkaisu. <https://urn.fi/URN:NBN:fi-fe2018052424593> (14.10.2022)
- Vellonen M., Konu A., Kivimäki H., Koivisto AM. & Joronen K. (2015) Lukiolaisten ja ammattiin opiskelevien liikunta- ja syömistottumusten yhteys ruutuaikaan. *Sosiaalilääketieteellinen aikakauslehti* **52**(1), 33–46.
- Välimäki M., Yang M., Vahlberg T., Lantta T., Pekurinen V., Anttila M. & Normand SL. (2019) Trends in the use of coercive measures in Finnish psychiatric hospitals: a register analysis of the past two decades. *BMC Psychiatry* **19**(230), 1–15.
- Välkkilä L., Joronen K., Koivisto AM. & Kanste O. (2020) Kouluterveydenhoitajan palvelut: sosiodemografisten tekijöiden yhteys nuorten kokemuksiin. *Sosiaalilääketieteellinen aikakauslehti* **57**(2), 155–167.
- Wargh M., Konu A., Kivimäki H., Koivisto AM. & Joronen K. (2015) 8.- ja 9.- luokkalaisten itsearvioitu masentuneisuus ja perhetekijät. *Sosiaalilääketieteellinen aikakauslehti* **52**(1), 78–89.

Katja Joronen, TtT, professori, Hoitotieteen laitos, Lääketieteellinen tiedekunta, 20014 Turun yliopisto, katja.joronen@utu.fi

Outi Kanste, TtT, dosentti, yliopistonlehtori, Lääketieteen tekniikan ja terveystieteiden tutkimusyksikkö, PL 5000, 90014 Oulun yliopisto, outi.kanste@oulu.fi

Kaisa Mishina, TtT, erikoistutkija, Kliininen laitos, Lastenpsykiatrian tutkimuskeskus, Lääketieteellinen tiedekunta, 20014 Turun yliopisto, kaisa.mishina@utu.fi