

Rekisteritietojen käyttö sosiaalitieteellisessä tutkimuksessa 2000-luvun alussa

ARI-PEKKA SIHVONEN – MIKA GISSLER

Johdanto

Suomen sosiaali- ja terveystilastojärjestelmä perustuu pääosin yksilötason rekistereihin. Tätä täydentävät mm. Kansaneläkelaitoksen ja Eläketurvakeskuksen mittavat hallinnolliset rekisterit sekä Tilastokeskuksen elinoloja kuvaavat rekisterit.

Terveysrekisterien historia alkoi jo 1950-luvulla, jolloin alettiin rekisteröidä mm. syöpäsairauksia (1953), psykiatrista hoitoa (1957) sekä uusia tuberkuloosi- ja sukupuolitauteja (1958). Rekisterien toimintaa edisti yksiselitteisten henkilötunnusten käyttöönotto vuosina 1964–1968. Tämä helpotti rekisterien kokoamista, paransi niiden laatua ja teki rekisterien yhdistämisen vaivattomammaksi. Pääosa edelleen toimivista terveysrekistereistä perustettiin 1980-luvun loppuun mennessä, kuten taulukko 1 osoittaa (Gissler & Haukka 2004).

Sosiaalihuollossa rekisteritietojen keruu alkoi verkkaisemmin. Eläkejärjestelmän kehityksen myötä Eläketurvakeskus aloitti eläketietojen keruun vuonna 1962 ja Kansaneläkelaitos eri sosiaalietuuksien keruun vuonna 1964. Muita kuin hallinnollisia sosiaalirekistereitä ovat ainoastaan Stakesin ylläpitämät 1980- ja 1990-luvulla perustetut toimeentulotukirekisteri, lastensuojelurekisteri ja sosiaalihuollon hoitoilmoitusrekisteri. Näiden rekisterien tiedot kerätään kuntien sosiaalitoimistoista sekä julkisista ja yksityisistä sosiaalihuollon laitoksista, kuten vanhainkodeista, palveluasumisyksiköistä ja kehitysvammlaitoksista.

Erillislainsäädäntö turvaa rekisterien keruun tilastointia varten ja niiden hyödyntämisen tutkimuksessa. Henkilötietolain (523/1999) 11. § kieltää arkaluonteisten tietojen käsittelyn. Tällaisina tietoina pidetään tietoja, jotka kuvaavat mm.

terveydentilaa, sairautta tai vammaisuutta taikka hoitotoimenpiteitä tai niihin verrattavia toimia ja sosiaalihuollon tarvetta tai sosiaalihuollon palveluja, tukitoimia ja muita sosiaalihuollon etuuksia. Saman lain 12. §:n mukaan kyseinen kieltö ei kuitenkaan estä tietojen käsittelyä historiallista tai tieteellistä tutkimusta taikka tilastointia varten. Kansallisella tasolla kerättävät terveysrekisterit on lueteltu terveydenhuollon valtakunnallisista henkilöreistereistä annetussa laissa (556/1989) ja asetuksessa (774/1989). Stakesin sosiaalirekisterien pito puolestaan perustuu Stakesista annettuun Tilastolakiin (409/2001).

Henkilötietolain lähtökohtana on, että tilastointi ja tutkimus tehdään ilman yksilötason tietoja aina, kun se on mahdollista. Henkilötietojen käsittelemiseen pyydetään siis ensisijaisesti tutkimusaineistoon kuuluvien henkilöiden suostumus, mikäli se on mahdollista. Jos suostumuksen hankkiminen ei kuitenkaan ole mahdollista esimerkiksi tietojen suuren määrän, tietojen iän tai muun sellaisen syyn vuoksi, tietoihin tutustumiseen voidaan hakea viranomaislupaa. Pääsääntöisesti tutkija tai tutkimusryhmä saa tutkimusluvan, vaikkei se aina ole alkuperäissuunnitelmien mukainen.

Runsaasta tutkimuskäytöstä huolimatta olemassa olevat rekisteritiedot ovat alikäytössä. Hallinnollisten rekisterien tutkimuskäyttöä edistääkseen Suomen epidemiologinen yhdistys ry. ja Sosiaalilääketieteen yhdistys ry. ehdottivat Suomen Akatemian terveyden tutkimuksen toimikunnalle rekisteritutkimusta tukevan keskuksen perustamista vuonna 2001. Esiselvityksen (Gissler & al. 2002) jälkeen Suomen Akatemian myönsi Rekisteritutkimuksen tukikeskukselle määräaikaisen infrastruktuurirahoituksen vuosille 2003–2005.

Tukikeskuksen tavoitteena on edistää kansallis-

Taulukko 1. Suomen terveys- ja sosiaalirekisterit

Vuosi	Rekisteri	Nykyinen pitäjä
Terveysrekisterit		
1953	Syöpärekisteri	Stakes (Suomen syöpärekisteri)
1955	Terveystieteiden tutkimuskeskusrekisteri	TEO
1956	Tuberkuloosisairaaloiden poistoilmoitukset	(osa Stakesin hoitoilmoitusrekisteriä)
1957	Psykiatristen sairaaloiden poistoilmoitukset	(osa Stakesin hoitoilmoitusrekisteriä)
1958	Uudet tuberkuloositapaukset	(osa KTL:n tartuntatautirekisteriä)
1958	Sukupuolitaudit	(osa KTL:n tartuntatautirekisteriä)
1960	Yleissairaaloiden poistoilmoitukset	(osa Stakesin hoitoilmoitusrekisteriä)
1963	Epämuodostumarekisteri	Stakes
1964	Ammattitautirekisteri	Työterveyslaitos
1966	Lääkkeiden haittavaikutusrekisteri	Lääkelaitos
1967	Sairaalan poistoilmoitukset	(osa Stakesin hoitoilmoitusrekisteriä)
1968	Joukkotarkastusrekisteri (kohdunkaulan- ja rintasyöpä)	Stakes (Suomen syöpärekisteri)
1969	Kuolemansyyrekisteri	Tilastokeskus
1977	Raskaudenkeskeytys- ja steriloisimisrekisteri	Stakes
1979	Syöpävaarallisille aineille altistuneiden rekisteri	Työterveyslaitos
1980	Implanttirekisteri (ortopediset implantit)	Lääkelaitos
1982	Huumeaineiden seurantarrekisteri	Lääkelaitos
1983	Näkövammaarekisteri	Stakes (Näkövammaisten keskusliitto)
1987	Syntymärekisteri	Stakes
1989	Tartuntatautirekisteri	Kansanterveyslaitos
1994	Implanttirekisteri (hammasimplantit)	Lääkelaitos
1994	Terveystieteiden tutkimuskeskusrekisteri	Stakes
Sosiaalirekisterit		
1962	Eläkerekisteri	Eläketurvakeskus
1964	Sairausvakuutuksen korvaukset ¹	Kansaneläkelaitos
1985	Toimeentulokirekisteri	Stakes
1991	Lastensuojelurekisteri	Stakes
1994	Sosiaalihuollon hoitoilmoitusrekisteri	Stakes
<p>¹Kansaneläkelaitoksen rekisterit kattavat vuodesta 1964 erityiskorvattavat lääkkeet, vuodesta 1970 kansaneläkkeet, perhe-eläkkeet ja lasten vammaistuet, vuodesta 1971 sairauspäivärahat ja sairausvakuutuksen korvaamat yksityisen terveydenhuollon käynnit, vuodesta 1978 kuntoutukset, vuodesta 1985 sotilasavustukset ja työttömyyskorvaukset, vuodesta 1993 perhe-etuisuudet ja lastenhoitoetuisuudet, vuodesta 1994 äitiysavustukset, työmarkkina-avustukset ja asumistuet, vuodesta 1996 reseptilääkkeet ja sairausvakuutuksen korvaamat yksityisen terveydenhuollon toimenpiteet sekä vuodesta 1997 opintotuet.</p> <p>TEO = Terveystieteiden tutkimuskeskus</p> <p>Lähde: Gissler & Haukka 2004</p>		

ten rekisterien tutkimuskäyttöä erityisesti terveys- ja sosiaalitieteissä

- tukemalla rekisteriaineistoihin perustuvan tutkimuksen suunnittelua ja toteutusta,
- parantamalla tutkijoiden rekisteriaineistojen käytön valmiuksia,
- lisäämällä eri rekisterien välistä yhteistyötä ja
- kehittämällä rekisteriaineistojen käyttöön liittyviä menettelytapoja.

Tavoitteidensa toteutumiseksi tukikeskus on järjestänyt seminaareja ja koulutustilaisuuksia, tiedottanut aineistojen käyttömahdollisuuksis-

ta, erilaisista tapahtumista ja koulutuksesta, tehnyt konkreettisia kuvauksia rekistereistä ja niiden tietosisällöistä, kerännyt yksityiskohtaisia käytännön ohjeita lupien hakemiseen sekä luonut rekisteritutkimusta tukevia viranomaisverkostoja ja sähköpostilistoja.

Tukikeskuksen toiminta painottui alkuvaiheissa terveysrekistereihin, mutta myöhemmin sitä on laajennettu kattamaan myös sosiaalirekisterit. Yhtenä osana sosiaalirekisterien hyödyntämistä tukevaa toimintaa Rekisteritutkimuksen tukikeskus teki selvityksen, jossa kuvataan, kuinka pal-

jon ja mitä rekisteriaineistoja sosiaali- ja yhteiskuntatieteellisessä tutkimuksessa käytetään. Selvityksen tarkoitus oli kartoittaa, mitä ongelmia rekistereiden käytössä koetaan olevan, ja etsiä ratkaisuja keskeisimpiin ongelmiin. Lisäksi kerättiin toiveita ja ehdotuksia siitä, miten rekisterien käyttöä sosiaali- ja yhteiskuntatieteellisessä tutkimuksessa parhaiten voisi edistää. Erikseen kerättiin bibliografia rekisteritietoja hyödyntävästä sosiaalitieteellisestä tutkimuksesta ja tutkimushankkeista. Se sisältää noin 350 julkaisun bibliografiset tiedot ja lisäksi joitain linkkejä Internetissä oleviin tutkimusryhmien, -yksiköiden ja -laitosten julkaisuluetteloihin. Sen saa käyttöönsä Rekisteritutkimuksen tukikeskuksen kotisivulta (<http://www.rekisteritutkimus.fi>), eikä sitä tässä analysoida tarkemmin.

Tietojenkeruu

Tietojen kerääminen tapahtui pääasiassa kyselylomakkeen avulla, aluksi ohjaamalla ihmisiä täyttämään lomake Internetissä (http://www.ktl.fi/portal/suomi/yhteistyoprojektit/retki/toiminta/tutkimus/kysely_rekisteriaineistot_sosiaalitieteissa/) ja myöhemmin jakamalla lomakkeen elektronista versiota sähköpostitse oletetuille rekisteritietoja käyttäville tai niiden käytöstä kiinnostuneille tutkijoille. Lomakkeen lisäksi selvityksestä vastannut tutkija (Ari-Pekka Sihvonen) otti yhteyttä joihinkin eri yliopistojen ja tutkimuslaitosten keskeisiin henkilöihin kartoittaakseen heidän näkemyksiään rekisteritietojen hyödyntämisestä omassa yksikössään, laitoksessaan, tiedekunnassaan ja tieteenalallaan.

Tietojenkeruun menetelmänä oli lumipallo-otanta eli vastaajia pyydettiin myös ehdottamaan, keihin muihin rekisteritietoja käyttäviin tai rekisteritietojen käytöstä kiinnostuneisiin tutkijoihin tai tutkijaryhmiin kannattaisi ottaa yhteyttä. Tietojen kerääminen aloitettiin lokakuussa 2004, ja se jatkui maaliskuun loppuun 2005. Suurin osa vastauksista saatiin alkuvuonna 2005. Kyselyä lähetettiin 11 yliopistoon ja korkeakouluun sekä 38 muuhun tutkimuslaitokseen (liite). Kyselyn painopiste oli (sektori)tutkimuslaitoksissa työskentelevien kokemuksissa rekisteritietojen käytöstä sosiaalitieteellisissä tutkimuksissa, ei niinkään rekisterejä hallinnoivissa, ylläpitävissä ja tuottavissa laitoksissa työskentelevien kokemuksissa.

Vastaajia oli kaikkiaan 114, joista 87 (77 %)

kertoi käyttäneensä rekisteritietoja tutkimukseensa. Tutkimuksen vastausprosenttia on vaikea arvioida, koska mukana on myös tutkimusryhmien ja (tutkimus)laitosten koottuja vastauksia. On myös mahdotonta tietää, monelleko henkilölle kysely on lopulta levinnyt, koska sitä on levitetty myös esimerkiksi eri tutkimusryhmien, laitosten ja tieteenalojen sisäisten sähköpostilistojen kautta. Arviomme mukaan vastaajia oli melkein kolmannes niistä, joille kysely lähetettiin. Vastauksia saatiin 8 yliopistosta ja korkeakoulusta sekä 24 muusta tutkimuslaitoksesta.

Tutkimuskäyttöä paljon

Vastaajat olivat käyttäneet rekisteritietoja sosiaali- ja yhteiskuntatieteellisessä tutkimuksessa suhteellisen paljon. Vastaajien raportoimat tutkimushankkeet kattoivat sosiaalitieteellistä tutkimuskenttää laajasti. Projektien sijoittaminen tieteenalan mukaan on tehty pääasiassa vastaajan oman näkemyksen mukaisesti. Ne kattavat seuraavat tieteenalat: epidemiologia, gerontologia, kansantaloustiede, kansanterveystiede, kriminologia, naistutkimus, nuorisotutkimus, oikeustaloustiede, oikeustiede, rikospsykologia, rikostiede, sosiaali- ja yhteiskuntapolitiikka, sosiaalipsykologia, sosiologia, terveystaloustiede, väestötiede ja yhteiskuntamaantiede. Viime kädessä luokittelu on tietysti kirjoittajien laatima ja viitteellinen. Useat tutkimushankkeet voitaisiin sijoittaa useammallekin kuin yhdelle tieteenalalle.

Kyselylomakkeessa mainittiin esimerkkinä keskeisimpiä sosiaali- ja terveysalan rekistereitä rekisterinpitäjän mukaan aakkosjärjestyksessä sekä avoin luokka muiden rekistereiden raportoimiseksi. Melkein kaikki vastaajat ovat käyttäneet ainakin yhtä Tilastokeskuksen rekisteriä (taulukko 2). Myös Väestörekisterikeskuksen, Kelan ja Stakesin rekistereitä käytetään paljon (noin 40 prosenttia vastaajista). Rikos- ja oikeusalan, työvoimaviranomaisten ja verottajan rekistereitä käyttäneitä on reilu neljännes.

Ongelmat ja vaikeudet

Rekisteritietoja käyttäneistä 87 vastaajasta melkein neljännes (24 %) ei raportoinut mainittavia rekisteritietojen käytön vaikeuksia. Tietojen tutkimuskäyttöä koskevan luvan epäminen oli suh-

teellisen harvinaista: yhdeksän vastaajaa ei ollut saanut lupaa haluamiensa rekisteritietojen käyttöön. Useimmiten ongelmaksi koettiin luvan saamisen hitaus tai lupaprosessiin liittyvät vaikeudet (noin kolmanneksella). Myös tekniset ongelmat rekisteritietojen käytössä ja epäselvyydet rekisterien tietosisällöstä vaikeuttivat rekisteritietojen käyttöä (melkein 30 %). Aineiston analyysissä ja menetelmällisissä kysymyksissä raportoitiin suhteellisen harvoin ongelmia. (Taulukko 3.)

Avovastauksien perusteella pystyimme erottamaan kuusi rekisteritietojen ja -aineistojen käytön ongelmakohtaa:

1) Rekisterien yleiset ongelmat: Rekisterin suunnittelussa ei oteta huomioon mahdollista myöhempää tutkimuskäyttöä. Tietoja poistetaan rekisteristä liian aikaisin. Rekisteritiedot laahaavat parisen vuotta nykypäivää jäljessä (erityisesti Tilastokeskus). Tiedon luovuttamisen teknisiä ratkaisuja ei ole mietitty. Rekistereiden tietosisällöt on huonosti dokumentoitu.

2) Rekisterien tietosisältöön liittyvät ongelmat: Tietoja ei saa toivomallaan tarkkuudella (esim. Tilastokeskuksen tiedot ammatista, asuinkunnasta, asuinalueesta, koulutuksesta, kielestä tai kansalaisuudesta). Toisaalta osasta aineistoista vain osa tiedoista on saatavissa (esim. Kelan sairaspäivärahan myöntöpäätösten syydiagnoosit on tallennettu vain 8. ja 28. päivä syntyneillä).

3) Tietojen laatuun ja aineistojen muodostamiseen liittyvät ongelmat: Rekisteriaineistoissa on havaittu virheitä ja väärinmuodostettuja muuttujia, erityisesti Tilastokeskuksen osalta. Rekisteritietojen laatu vaihtelee mm. tiedontuottajien erilaisten raportointikäytäntöjen vuoksi (esim. Stakesin hoitoilmoitusrekisteri). Eri vuosien koodaukset eivät ole yhdenmukaisia.

4) Luvanhakuprosessin pituus ja aineiston käyttöön saamisen hitaus (erityisesti Tilastokeskus).

”Tietosuojaan mukainen luvan hakuprosessi on kyllä hyvin hoidettu ja selkeä, mutta kankea. Se edellyttää, että jo ennen tutkimusta pystytään määrittelemään tarkasti se, mitä tietoja käytetään hyväksi. Yhtä rekisteriä käytettäessä tämä ei ole iso ongelma, mutta jos useita rekistereitä joudutaan yhdistelemään, joudutaan tutkimusta suuntaamaan tietojen yhteensopivuuden ja saatavuuden mukaan. Jos tämä vaatii uutta lupaa, sen hakuprosessi voi romuttaa projektin aikataulun. Esimerkiksi yhteensä 6 kuukauden projektissa ei ehditä hakea uutta lupaa. Tälle tarttis varmaan tehdä jotain.” (Vastaaja tutkimuslaitoksesta)

5) Rekisteriaineistojen kalleus (erityisesti Tilastokeskus ja Kansaneläkelaitos). Tutkijat kertoivat

Taulukko 2. Kyselyssä mainittuja rekisterien päättäjiä ja rekistereitä (N = 87)

Lomakkeessa mainitut rekisterit	Mainintoja
Tilastokeskus	71
Väestörekisterikeskus	35
Stakes	34
Kansaneläkelaitos	32
Verotusviranomaiset	24
Työvoimaviranomaiset	22
Oikeusviranomaiset	22
Eläketurvakeskus	18
Kunnan rekisterit ¹	8
Työterveyslaitos	3
Kansanterveyslaitos	2
Puolustusvoimat	1
Säteilyturvakeskus	1
Muut ²	15

¹Esimerkiksi Helsingin kaupungin henkilöstörekisteri, Helsingin tietokeskuksen kokoomarekisterit, Vantaan kaupungin rekisterit, sosiaali- ja terveystoimen maksupäätökset.

²Eläkelaitokset, ulkomaiset rekisterilähteet (vertailututkimuksissa), maa- ja metsätalousministeriön maaseutuelinkeinorekisteri, paikallisen poliisin rekisteri, kauppa- ja teollisuusministeriön ja Tekesin yritystutkimuksia koskevia päätöksiä sisältävät tietokannat, sairaaloiden ja vanhainkotien rekisterit, tutkintorekisteri, yhteishakurekisteri ja ulkoministeriön rekisteri lähetystöissä myönnettyistä työluvista.

Taulukko 3. Mitä vaikeuksia tai ongelmia vastaajilla on ollut rekisteritietojen käytössä?

Ei ole ollut ongelmia (21)
 Vaikeus saada lupaa tai luvan saannin hitaus (28)
 Tekniset ongelmat rekisteritietojen käytössä tai yhdistelyssä (26)
 Epäselvyys siitä, mitä tietoja rekisterit sisältävät (24)
 Ei saanut lupaa rekisterien käyttöön (9)
 Aineiston analyysit ja menetelmät (7)
 Muita ongelmia (24)

Vastaajia oli 87 ja ongelmia 119.

myös saaneensa luvan hankintavaiheissa väärää tietoa arvonlisäverosta, jota ei ollut budjetoitu.

”Olen käsittänyt, että rekisteritiedot ovat erittäin kalta, joten en ole käyttänyt niitä.” (Vastaaja yliopistosta)

6) Vaikeudet aineiston muokkaamisessa, käytössä ja analyyseissä: Erityisinä ongelmina mainittiin menetelmien osaamisen vähäisyys, it-tuen puute ja konekapasiteetin niukkuus. Helsingin ulkopuolella työskenteleville ongelmaksi muodostui

se, ettei aineistoa tietosuojasyistä luovutettu rekistereitä pitävän organisaation (Tilastokeskus) ulkopuolelle. Aineistokulujen lisäksi ylimääräisiä kustannuksia syntyi matka- ja oleskelukustannuksista.

Jotkut vastaajat raportoivat, ettei aineistoja saanut enää käyttöön työsuhteen päätyttyä. Tämä viittaa siihen, ettei tutkijoilla ole tarpeeksi tietoa aineiston käyttöluvan hakemisesta ja siihen liittyvistä säädöksistä ja käytännöistä. Tutkimusluvut myönnetään pääsääntöisesti laitoksille, joihin aineistojen käyttäjien on oltava työ-, opinto- tai vastaavassa suhteessa.

Vastaajia oli suhteellisen vähän suhteessa laitosten määrään, joten alueellisista tai paikallisista erityisongelmista rekisteritietojen käytön osalta on vaikea sanoa mitään. Yleinen huoli kvantitatiivisten menetelmien opettamisesta ja osaamisesta sosiaalitieteissä heijastuu osasta vastauksia, mutta ei kovinkaan voimakkaasti. Tieteenalojen välillä ei juuri ollut eroja ongelmista tai niiden määrässä. Aineistosta kuitenkin erottuu kriminologien huoli rekisteritietojen ja -aineistojen tutkimuskäyttöön saamisen vaikeudesta

Miten rekisteritutkimusta voisi tukea?

Osittain vastaajien ehdottamat toimet rekisteriaineistojen käytön edistämiseksi liittyivät edellä kuvattuihin, koettuihin rekisteriaineistojen käytön ongelmiin. Moni ehdotus kattoi sen toiminnan, jota Rekisteritutkimuksen tukikeskus jo tekee (ks. johdanto). Jaoimme toiveet seitsemään luokkaan:

1) Vastaajat toivoivat kanavia kokemusten vaihtamiseen: lisää seminaareja ja koulutustilaisuuksia.

2) Tiedotuksen lisääminen, erityisesti opiskelijoille ja aloittaville tutkijoille, katsottiin tärkeäksi.

3) Rekisteritutkimus hyötyisi koulutuksen lisäämisestä, erityisesti menetelmien osalta.

4) Olisi hyödyllistä jakaa vieläkin konkreettisempaa tietoa eri rekistereistä ja niiden käyttömahdollisuuksista.

5) Tukiverkoston luominen ja ylläpitäminen voisivat parantaa yhteistyötä viranomaisten ja tutkijoiden välillä.

6) Eri laitosten ja tutkimusryhmien välisen sekä (tutkimus)laitosten ja yliopistojen välisen yhteistyön tukeminen edistäisi rekisteritutkimusta. Esimerkkeinä mainittiin aineistojen yhteishankinnat

ja olemassa olevien aineistojen jakaminen eri tutkimusryhmien välillä.

7) Yleinen rekisteritutkimusalan ”edunvalvonta” on tärkeää, sillä vastaajien mielestä rekisteritietojen olisi oltava halpoja tai kokonaan maksuttomia.

”PALJON panostusta ja apua viranomaisten vakuuteluun, että rekisterit todella aukeaisivat tutkimuskäyttöön, sillä niin monet hyvät hankkeet tyssäävät viranomaisten vitkasteluun ja kielteiseen kantaan rekistereiden käyttämisestä tutkimuskäyttöön.” (Vastaaja tutkimuslaitoksesta)

Johtopäätökset

Rekisteritutkimuksen tukikeskus on toiminut kaksi vuotta. Vuoden 2005 alussa tehdyn arviointiraportin (Rekisteritutkimuksen ..., 2005) mukaan sen toiminta on lähtenyt hyvin käyntiin, varsinkin kun otetaan huomioon tukikeskuksen rajalliset voimavarat. Selvitys osoitti, että tukikeskuksen toiminnan tunnetuksi tekemistä on jatkettava. On kuitenkin oletettava, että terveystutkijoiden piirissä Rekisteritutkimuksen tukikeskus on tullut tutuksi aiemmin kuin sosiaali- ja yhteiskuntatieteiden puolella.

Selvityksen pohjalta toiveet rekisteritutkimuksen edistämistoimista voidaan tiivistää seuraavasti: Rekisteritutkimuksen tukikeskuksen tähänastisen toiminnan yleiseen linjaan oltiin tyytyväisiä, tosin jo vakiintuneen toiminnan lisäämistä ja hienosäätöä esitettiin. Erityisesti ajankohtaisten aiheiden ja rekisterien käyttömahdollisuuksista valistamista ja propagoimista pidettiin tärkeänä.

Vastaajat toivoivat lisää ja vielä tarkempia kuvauksia tukikeskuksen Internet-sivuille eri rekistereistä ja niiden tietosisällöistä, esimerkiksi kriminologisen ja taloustieteellisen tutkimuksen kannalta hyödyllisistä rekistereistä. Lisäksi toivottiin, että hallinnolliset rekisterinpitäjät vielä parantaisivat omaa dokumentointiaan. Koulutuskysyntään – esimerkiksi rekisteriaineistojen yleiskuvausten, menetelmien ja luvan hakemisen osalta – vastaa osaltaan vuoden 2005 kuluessa julkaistava rekisteritutkimusta koskeva verkkokurssimateriaali. Se on valmistuttuaan opiskelijoiden, opettajien ja tutkijoiden vapaassa käytössä Suomen virtuaaliyliopiston portaalin kautta (<http://www.virtuaaliyliopisto.fi>).

Rekisteriviranomaisten ja tutkimuslaitosten yhdyshenkilöiden ja asiantuntijoiden verkostoa

on laajennettava, ja sitä olisi hyödynnettävä rekisteritutkimusta koskevan koulutuksen ja muiden tapahtumien tiedottamisessa. Yhdyshenkilöt voisivat myös toimia hyvänä välikätenä tutkijoiden ja rekisterinpitäjien välillä, kerätä tietoa ongelmista ja välittää niistä tietoa eteenpäin sekä rekisteristä vastaaville viranomaisille että tukikeskukselle. He voisivat myös koota kokemuksia ja aineistoa siitä, miten tutkimuskäyttö on nosta- nut esiin rekisterien korjattavissa olevia puutteita ja kohentanut niiden laatua sekä hallinnossa että tutkimuksessa. Tämä todennäköisesti parantaisi myös rekisteriviranomaisten motivaatiota tutkimusyhteistyöhön. Lisäksi verkostot voivat edistää tutkimusnäkökulman huomioon ottamista rekisterien uudistamisessa ja uusien tietojärjestelmien suunnittelussa.

Eri toimijoiden hallinnollisten ja tieteenalojen välisten rajojen ylittävän yhteistyön tukeminen katsottiin tärkeäksi. Erityisesti mainittiin yhteisten, uusien tutkimusaineistojen hankkiminen ja käyttö sekä mahdollisuus jakaa olemassa olevia aineistoja eri tutkijaryhmien välillä. Yksi tapa edistää rekisteriaineistojen yleistä saatavuutta on anonyymien tutkimusaineistojen arkistointi Yhteiskuntatieteelliseen tietoarkistoon (FSD), kuten erilaisten tunnisteettomien survey-aineistojen osalta tehdään. Rekistereihin perustuvien tutkimusaineistojen osalta olisi neuvoteltava FSD:n, tietosuojavaltuutetun toimiston ja tärkeimpien tutkimuslupien myöntäjien kanssa toimivat pelisäännöt, jotta tällainen vapaa arkistointi mahdollistuisi.

Moni vastaaja korosti yleisellä tasolla rekisteritutkimuksen edunvalvonnan tärkeyttä. Open access -periaatteet saivat laajaa kannatusta. Rekisteritutkimuksen tukikeskuksen olisi yhdessä tiedeyhteisön kanssa vaikutettava esimerkiksi virastoihin, eduskuntaan ja ministeriöihin ym. rekisteriaineistojen saamiseksi edullisemmin tutkimuskäyttöön. Muissa Pohjoismaissa rekisteritutkimusta on edistetty mm. jakamalla rekisteritutkimukseen suunnattuja apurahoja (Ruotsi) ja

TIIVISTELMÄ

Ari-Pekka Sibvonen & Mika Gissler: Rekisteritietojen käyttö sosiaalitieteellisessä tutkimuksessa 2000-luvun alussa

Rekisteritutkimuksen tukikeskus perustettiin vuonna 2003 tukemaan hallinnollisten rekisterien tutkimuskäyttöä. Lokakuun 2004 ja maaliskuun 2005 välillä tukikeskus teki selvityksen rekisteriaineistojen käy-

myöntämällä rekisteriviranomaisille määrärahoja hallinnollisten rekisterien tutkimuskäytön edistämiseksi (Tanska).

Suomessa on esitetty valtion maksuperustelain ja -asetuksen uudistamisen yhteydessä, että viranomaistietojen tutkimuskäytölle myönnettäisiin erikoisasema käyttäjämaksujen pienentämiseksi (Sosiaali- ja terveydenhuollon ..., 2003). Viranomaisten näin menettämät tulot korvattaisiin erillisellä määrärahalla tai myöntämällä lisävirkoja merkittävimmille rekisterinpitäjille aineistopaimintojen nopeuttamiseksi ja rekisteri-infrastruktuurin parantamiseksi. Asia on kuitenkin edennyt kovin hitaasti.

Esille nostetuista pulmista huolimatta rekisteritutkimusta ja sen edistämistä pidettiin tärkeänä. Olemassa olevien aineistojen käyttö on taloudellista, ja se avaa uusia mahdollisuuksia uusille tutkimuskysymyksille ja -asetelmille. Tätä edistää myös esimerkiksi sosiaali- ja terveydenhuollon avopalveluiden, hoitoketjujen ja kustannustietojen laajentuva rekisteröinti. Kehittyvä it-teknologia puolestaan nopeuttaa rekisteriaineistojen valmistumista ja parantaa perusrekisterien laatua, joten tutkijat voivat saada tiedot käyttöönsä nopeammin ja luotettavammin. Suurin osa näistä kehitysehdotuksista ja tulevaisuuden näkymistä ei tietenkään koske vain sosiaali- ja yhteiskuntatieteellistä tutkimusta, vaan ne soveltuvat yhtä hyvin terveystieteisiin.

Rekisteritutkimuksen uhkia puolestaan ovat tiukentuva tietosuojalainsäädäntö tai sen tulkinta joko kansallisesti tai EU-tasolla, vähenevät määrärahat rekisterien ylläpitoon, rekisteritietojen tutkimuskäytön kallistuminen ja yleisen tutkimusrahoituksen väheneminen (Gissler 1999). Näitä vastaan on taisteltava eri tasoilla ja pyrittävä vaikuttamaan erityisesti poliittisiin päättäjiin ja viranomaisiin. Tässä työssä Rekisteritutkimuksen tukikeskuksella on selvä roolinsa, mutta sen rinnalla tarvitaan yksittäisten tutkijoiden, tutkimusryhmien ja -laitosten sekä tieteellisten seurojen ja yhteisöjen panosta.

töstä sosiaali- ja yhteiskuntatieteellisessä tutkimuksessa. Tiedot kerättiin Internetissä olevan kyselylomakkeen avulla ja jakamalla lomaketta sähköpostitse. Vastaajia oli 114, yhteensä 32 korkeakoulusta tai tutkimuslaitoksesta. Vastaajat raportoivat runsaasti rekisteritietojen käyttöä, eniten Tilastokeskuksen, Väestörekisterikeskuksen, Kelan ja Stakesin rekistereiden hyödyntämistä. Neljännes vastaajista ei raportoinut mai-

nittavia rekisteritutkimuksen vaikeuksia. Ongelmat liittyivät rekisterien toiminnan ja tietosisällön ongelmiin, tietojen laatuun ja aineistojen muodostamiseen, luvanhakuprosessin pituuteen ja tietojen saannin hitauteen, tietojen kalleuteen sekä aineiston muokkaamisen, käytön ja analyysin vaikeuksiin. Vastaajat toivoivat lisää rekisteritutkimusta koskevia seminaareja ja koulutustilaisuuksia, tiedotuksen ja koulutuksen lisäämistä,

viranomaisten ja tutkijoiden välisen yhteistyön lisäämistä ja tutkimusryhmien välisen yhteistyön tukemista. Rekisteritutkimuksen uhkina pidettiin tiukentuvaa tietosuojalainsäädäntöä ja tutkimusrahoituksen vähenemistä. Sen vuoksi Rekisteritutkimuksen tukikeskuskella olisi oltava selkeä rooli rekisteritutkimuksen edunvalvonnassa, jotta rekisteritietoja voitaisiin käyttää tutkimuksessa kohtuullisin kustannuksin.

KIRJALLISUUS

- GISSLER, MIKA: Hot mot registerforskning. Sosiaalilääketieteellinen aikakauslehti 36 (1999), 183–184
- GISSLER, MIKA & HAUKKA, JARI: Finnish health and social welfare registers in epidemiological research. Norsk Epidemiologi 14 (2004), 113–120
- GISSLER, MIKA & RIIHIMÄKI, HILKKA & KOSKELA, RIITTA-SISKO: Kansa kuunteli Retua: Raportti Rekisteritutkimuksen tukikeskusta koskevasta kuulemis-

laisuudesta 27.5.2002 Suomen Akatemiassa. Sosiaalilääketieteellinen aikakauslehti 39 (2002), 236–239

REKISTERITUTKIMUKSEN TUKIKESKUKSEN ARVIOINTIRAPORTTI. Julkaisematon raportti. Rekisteritutkimuksen tukikeskus, 2005

SOSIAALI- JA TERVEYDENHUOLLON TIETOUDISTUS 2005 -TYÖRYHMÄN RAPORTTI. Työryhmämuistioita 2003: 37. Helsinki: Sosiaali- ja terveysministeriö, 2003.

Liite. Kyselyn kattavuus

- Helsingin yliopisto: matematiikan ja tilastotieteen, sosiaalipsykologian, sosiologian, yhteiskuntapolitiikan, yleisen valtio-opin, kasvatustieteen, maantieteen ja kansanterveystieteen laitokset sekä Svenska social- och kommunalhögskolan
- Joensuun yliopisto: maantieteen, sosiologian, taloustieteiden, tilastotieteen, yhteiskuntapolitiikan ja kasvatustieteen laitokset
- Jyväskylän yliopisto: psykologian ja yhteiskuntatieteiden, filosofian (filosofia, sosiaalityö, sosiologia, valtio-oppi, yhteiskuntapolitiikka), matematiikan ja tilastotieteen sekä terveystieteiden laitokset sekä koulutuksen tutkimuslaitos
- Kuopion yliopisto: sosiaalipsykologian ja sosiologian, sosiaalipolitiikan ja sosiaalitalouden, sosiaalityön ja sosiaalipedagogiikan sekä terveyshallinnon ja -talouden laitokset
- Lapin yliopisto: yhteiskuntatutkimuksen ja sosiaalityön sekä menetelmätieteiden laitos
- Oulun yliopisto: kansanterveystieteen ja yleislääketieteen laitos
- Tampereen yliopisto: politiikan tutkimuksen, psykologian, sosiaalipolitiikan ja sosiaalityön, sosiologian ja sosiaalipsykologian, yhdyskuntatieteiden, oikeustieteiden, matematiikan, tilastotieteen ja filosofian, terveystieteiden ja taloustieteiden laitokset, Yhteiskuntatieteiden tutkimuslaitos, Työelämän tutkimuskeskus, Tampereen yliopiston kauppakorkeakoulu sekä kauppa- ja hallintotieteiden tiedekunta
- Turun yliopisto: koulutus sosiologian, sosiaalipolitiikan, sosiologian, taloustieteen, tilastotieteen ja valtio-opin laitokset sekä oikeustieteellinen tiedekunta
- Åbo Akademi: institutioner för samhällsekonomi och statistik, sociologi, statskunskap samt samhällsstatistik och statistik
- Vaasan yliopisto: markkinoinnin laitos
- Turun kauppakorkeakoulu: markkinoinnin ja talous-sosiologian laitokset

- Kansaneläkelaitos KELA
- Kansanterveyslaitos KTL
- Stakes
- Tilastokeskus
- Työterveyslaitos
- Kuntoutussäätiö
- UKK-instituutti
- Suomen gerontologian tutkimuskeskus
- Siirtolaisuusinstituutti
- Väestöliiton Väestöntutkimuslaitos
- Helsingin kaupungin tietokeskus
- Espoon kaupungin tilasto- ja tietopalveluyksikkö
- Vantaan kaupungin tilasto- ja tietopalveluyksikkö
- Opetushallitus
- Suomen Kuntaliitto
- Eläketurvakeskus
- Kuntien eläkevakuutus
- Elinkeinoelämän tutkimuslaitos ETLA
- Palkansaajien tutkimuslaitos
- Pellervon taloudellinen tutkimuslaitos PTT
- Valtion taloudellinen tutkimuskeskus VATT
- Oikeuspoliittinen tutkimuslaitos OPTULA
- Oikeusministeriön vankeinhoitolaitoksen rikosseuraamusvirasto RISE
- The European Institute for Crime Prevention and Control HEUNI, kriminaalipoliittinen osasto
- Vankeinhoidon koulutuskeskus
- Oikeusrekisterikeskus
- Poliisiammattikorkeakoulu PAKK
- Keskusrikospoliisi
- Kriminaalihuollon tukisäätiö
- Diakonia-ammattikorkeakoulu
- A-klinikkasäätiö
- Alkoholitutkimussäätiö
- Nuorisotutkimusseura ja Nuorisotutkimusverkosto
- Valtion nuorisosiain neuvottelukunta NUORA
- Kuluttajatutkimuskeskus
- Kulutustutkimuksen seura
- Karjalan tutkimuslaitoksen yhteiskuntatieteellinen osasto