

Tarja Laitinen, Arho Virkki ja Kimmo Porkka

Kansainvälisen harmonisointistandardin käyttöönotto on keskeistä suomalaisten terveystietojen tehokkaalle yhteiskäytölle

## FinOMOP: terveystietojen kansainvälinen harmonisointi

**A**rkipäivän potilastiedon (real-world data, RWD) ja siitä johdetun tutkimusnäytön (real-world evidence, RWE) tarve ja merkitys lisääntyy terveystieteellisessä tutkimuksessa ja tietojohdatusessa. COVID-19-pandemia osoitti, miten tärkeää lähes reaaliaikainen potilastietojen jakaminen ja yhdistäminen on, mutta myös sen, miten vaativaa se tällä hetkellä on suomalaisessa terveydenhuollossa. Olemme yhä eri toimittajien, paikallisten ratkaisujen ja yhteensopimattomien järjestelmien verkostossa, mikä on rajallisten resurssien haaskausta. Lisäksi Euroopan unionin tietosuojasetuksen (General Data Protection Regulation, GDPR) sekä etenkin sen suomalaisen implementoinnin, lain sosiaali- ja terveystietojen toisiokäytöstä (552/2019, toisiolaki) myötä eri rekisterinpitäjien tietoja yhdistävä tutkimus on lähes tyrehänyt. Syynä ovat lain epäselvyydet ja erilaiset paikalliset tulkinnat (1).

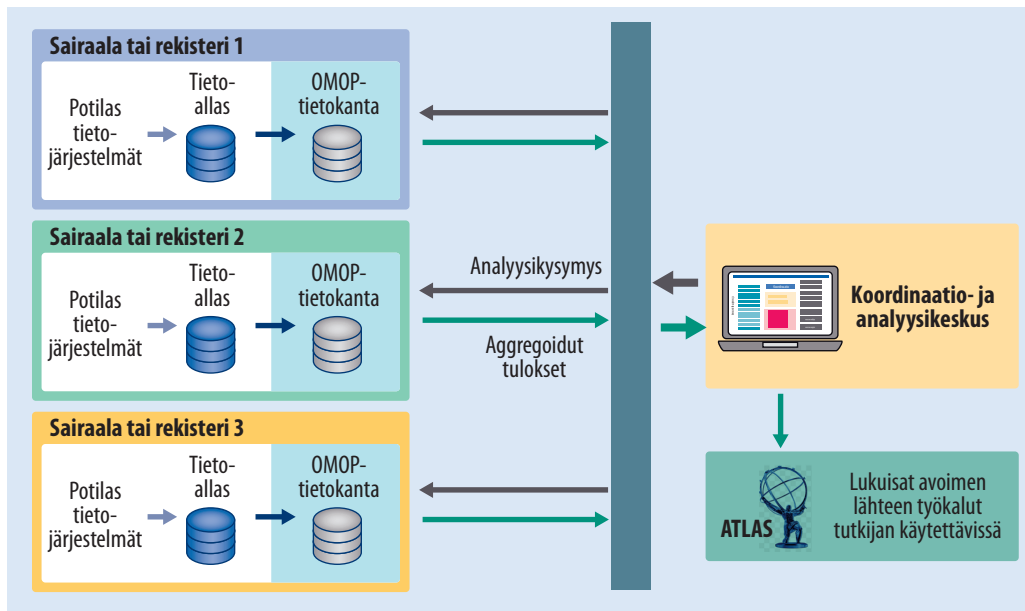
Satunnaistetut etenevät kliiniset tutkimukset (RCT) ja niistä tehdyt meta-analyysit ovat olleet lääketieteen kultastandardi vuosikymmenien ajan. Monen sairauden patogeneesin paremman tuntemuksen kautta hoidot ovat muuttumassa yhä enemmän potilaskohtaisesti yksilöllistettyyn suuntaan. RCT-asetelma soveltuu tämääntyyppisten hoitojen arviointiin huonosti, koska hoidon kohderyhmät voivat segmentoitua hyvin pieniksi.

RWD-tutkimukset ovat nopeampia ja edullisempia RCT-tutkimuksiin verrattuna sekä

pystyvät vastaamaan huolellisesti suoritettuna moniin erilaisiin kysymyksiin ja täydentämään RCT-tutkimuksen havaintoja (2–4). RWD-tutkimuksen luotettavuuden parantumisesta kertoo muun muassa lääkeviranomaisten muuttunut suhtautuminen niihin, ja myös prekliinisessä lääketutkimuksessa tunnustetaan jo erilaisia hybridimalleja (5). Luotettavuus on parantunut myös uusien tilastolisten analyysimenetelmien ansiosta, joiden käytöstä lisääntynyt laskentakapasiteetti on tehnyt mahdollista. Uudet menetelmät antavat täysin uusia mahdollisuuksia hallita puuttuvan tiedon ongelmaa sekä tarkastella ilmiöitä myös ilman etukäteishypoteesia.

Kehitystä on hidastanut potilastiedon yhteismitallisuuden ja yhteensopimattomuuden ongelma. Suomessa käytetään kansallisesti yhtenäisiä koodistoja, joita ylläpidetään Terveyden ja hyvinvoinnin laitoksen (THL) koodistopalvelimella (6). Kaikki koodistomme eivät kuitenkaan ole kansainvälisiä, ja kaikkiin koodistoihin syntyy helposti paikallisia vivahteita: diagnostiikan ja hoidon termejä voidaan käyttää täsmällisemmin tai löyhemmin, ja uusia termejä voidaan keksiä tarvelähtöisesti. Lääketieteen saralla on kehitteillä standardoituja kansainvälisiä tietomalleja, jotka eivät kuitenkaan ole vielä laajassa käytössä. Tällä hetkellä tietokantojen rakenteet vaihtelevat huomattavasti jo Suomen eri terveydenhuolto-organisaatioiden tietojärjestelmien välillä.

Sähköiset potilastiedot ja niitä hyödyntävät



**KUVA.** Hajautettu eli federoitu analyysimalli onnistuu sujuvasti vain, jos kaikki tutkimuskeskukset organisoivat potilastiedon samaan kansainväliseen koodistostandardiin ja tietomalliin.

OMOP = Observational Medical Outcomes Partnership

tutkimukset ovat tuoneet mukanaan monia uusia haasteita. Terveydenhuollon organisaatioiden ja tutkijoiden yhteinen velvoite on tietosuojan varmistaminen, mikä on synnyttänyt merkittävän määrän uutta lainsäädäntöä. Tutkimus on, kuten kuuluukin, kansainvälistä, mutta kansallinen lainsäädäntö tekee laajojen, maan rajat ylittävien potilastietorekistereiden muodostamisen vaikeaksi myös EU:n sisällä. Tämä on synnyttänyt idean hajautetuista (federoiduista) analyyseista, joka tarkoittaa sitä, että kliininen potilastieto on harmonisoitu ja tallennettu samanlaiseen tietomalliin kaikissa keskuksissa (**KUVA**). Kun tieto on yhteismittalista ja organisoitu samalla tavalla, voivat tutkimuskeskukset analysoida aineiston paikallisesti yhteisesti jaetulla tutkimusalgoritmilla, jonka jälkeen aggregaattitason tulokset voidaan yhdistää ja yhdessä asetettuun tutkimuskysymyksen voidaan vastata koko aineiston pohjalta. Tästä on se hyöty, että yksilötason potilastietoa ei siirretä keskuksien välillä, vaan ainoastaan johdetut tilastolliset tulokset siirtyvät.

Havainnointitutkimuskäyttöön erityisesti kehitetty OHDSI-yhteisön (Observational Health Data Sciences and Informatics) luoma

OMOP (Observational Medical Outcomes Partnership) -tietomalli on viime vuosina muodostunut laajalti tuetuksi sairaalatietojen harmonisoinnin standardiksi sekä Euroopassa että Yhdysvalloissa (7). EU:n IMI2-ohjelman EHDEN-hanke (European Health Data & Evidence Network) (8) on mahdollistanut OMOP-tietomallin laajan käyttöönoton Euroopassa, ja tällä hetkellä sillä on 143 datakumppania 27 eri maassa.

OMOP-standardi perustuu useisiin kansainvälisiin koodistoihin, kuten SNOMED CT, LOINC ja RxNorm. Useimmat Suomen koodistot perustuvat kansainvälisiin koodistoihin, ja niiden kääntäminen OMOP-standardiin on suhteellisen helppoa. Tutkimuksen lisäksi OHDSI-yhteisö ohjaa avoimen lähdekoodin ohjelmistokehitystä kliinisen potilastiedon standardoituun analysointiin, visualisointiin sekä laadun seurantaan. Kaikki nämä ohjelmat ovat vapaan lähdekoodin piirissä ja siten ilmaiseksi käytettävissä. Tällaiselle – tuhansien tutkijoiden jatkuvalle ohjelmistokehitykselle – on vaikeaa löytää kilpailijaa esimerkiksi yksittäisestä liikeyrityksestä. Toisaalta ohjelmistojen käyttöönotto yksittäisessä sairaalassa vaatii

enemmän asiaan paneutumista kuin mihin tavanomaisesti on ehkä totuttu.

Suomessa käynnistyi vuonna 2021 tutkijoiden yhteistyö FinOMOP-sanaston ja tietomallin luomiseksi, ja siihen osallistuvat kaikki yliopistosairaalat, FinnGen-tutkimus (FIMM, Helsingin yliopisto) ja THL. Työtä on rahoitettu EHDEN-hankkeina, ja mukana on myös ollut suomalaisia pieniä ja keskiuuria yrityksiä, jotka ovat saaneet EHDEN-sertifikaation. Yhteistyöllä varmistamme, että kaikkien toimijoiden OMOP-sanasto on identtinen. Ensimmäinen versio FinOMOP-sanastosta on tehty, ja



**TARJA LAITINEN, tutkimusjohtaja, molekyyli lääketieteen dosentti**  
Pirkanmaan sairaanhoitopiiri

**ARHO VIRKKI, tietopalvelujohtaja, lääketieteellisen matematiikan dosentti**

Tyks, Auria Tietopalvelu ja Turun yliopisto

**KIMMO PORKKA, ylilääkäri, professori**

HUS, Syöpäkeskus, tietohallinto ja Helsingin yliopisto

useissa yliopistosairaloissa OMOP-tietokanta on jo toimintavalmis. Toiminnan vakiinnuttaminen vaatii pysyvän resursoinnin.

Suomalaisten terveystietojen OMOP-harmonisointi mahdollistaa tietoturvallisen, hajutetun analyysin ilman, että potilaskohtaisia tietoja täytyy yhdistää, jolloin tietoturvaan ja lainsäädäntöön liittyvät ongelmat voidaan välttää. Se avaa Suomen korkealaatuiset tietolähteet uudella tavalla kansainvälisiin yhteistyöhankkeisiin, niin akateemisiin kuin kaupallisiinkin. ■

#### SIDONNAISUUDET

**Tarja Laitinen:** Muut sidonnaisuudet (Biopankkiosuuskunnan hallitus)

**Arho Virkki:** Luentopalkkio/asiantuntijapalkkio (Medaffcon), luottamustoimet (SOTE-tietojen toisilain ohjausryhmä), hankkeet (Sitra Isaacus, Toisiotietojen lupaviranomaisen toiminnan käynnistäminen)

**Kimmo Porkka:** Apuraha (BMS, Celgene, Incyte, Novartis, Pfizer), luontopalkkio/asiantuntijapalkkio (BMS, Celgene, Novartis, Pfizer), luottamustoimet (Euroopan hematologiyhdistys (EHA), hematologien biopankki FHRB)

#### KIRJALLISUUTTA

1. Bützow A. Selvitys tutkimusta koskevien sosiaali- ja terveydenhuollon säädösten vaikutuksesta tutkimuksen vapauteen ja tutkimus-, kehittämis- ja innovaatiotoimintaan. Opetus- ja kulttuuriministeriön julkaisuja 2021:42.
2. Ramagopalan SV, Simpson A, Sammon C. Can real-world data really replace randomised clinical trials? *BMC Med* 2020;18:13.
3. Eichler HG, Pignatti F, Schwarzer-Daum B, ym. Randomized controlled trials versus real world evidence: neither magic nor myth. *Clin Pharmacol Ther* 2021;109:1212–8.
4. Collins R. The magic of randomization versus the myth of real world evidence. *N Engl J Med* 2020;382:674–8.
5. Data analysis and real world interrogation network (DARWIN EU). Amsterdam: European Medicines Agency. <https://ema.europa.eu/en/about-us/how-we-work/big-data/data-analysis-real-world-interrogation-network-darwin-eu>.
6. Tiedonhallinta sosiaali- ja terveysalalla. Koodistopalvelu. Helsinki: Terveyden ja hyvinvoinnin laitos 3.12.2020. <https://thl.fi/fi/web/tiedonhallinta-sosiaali-ja-terveysalalla/koodistopalvelu>.
7. Recommendations for data standards in health data research. UK Health Data Research Alliance 2021. <https://ukhealthdata.org/wp-content/uploads/2021/12/211124-White-Paper-Recommendations-of-Data-Standards-v2-1.pdf>.
8. EHDEN: The European Health Data & Evidence Network Portal [verkko-tietokanta]. <https://portal.ehden.eu/>