



KÄYTÄNNÖN ESIMERKKEJÄ SATUNNAISTETTUIJEN KONTROLLOITUIJEN KOKEIDEN JÄLJITTELYSTÄ REAALIMAAILMAN AINEISTOISSA

JUSSI LEINONEN, BAYER OY

AJATELTAVAA

- Datat laatu → missä kohtaa RWD ei enää riitä RCT-jäljittelyyn?
- Missä kohtaa lääkekehityspolkua RWD–RCT-verailu on hyödyllinen?

MIKSI SATUNNAISTETUT KONTROLLOIDUT KOKEET (RCT:T)?

- Satunnaistus tasapainottaa tunnetut ja tuntemattomat sekoittavat tekijät.
- Mahdollistaa hoitovaikutusten kausaalisen tulkinnan, ei vain assosiaatioita.
- RCT:t ovat usein kalliita, hitaita ja joskus eettisesti tai logistisesti mahdottomia.
- Potilasjoukko voi olla valikoitunut, mikä rajoittaa yleistettävyyttä reaalimaailmaan.
- Tarvitaan täydentävää näyttöä reaalimaailman aineistoista.

MIKSI JÄLJITELLÄ RCT:TÄ REAALIMAILMAN AINEISTOLLA?

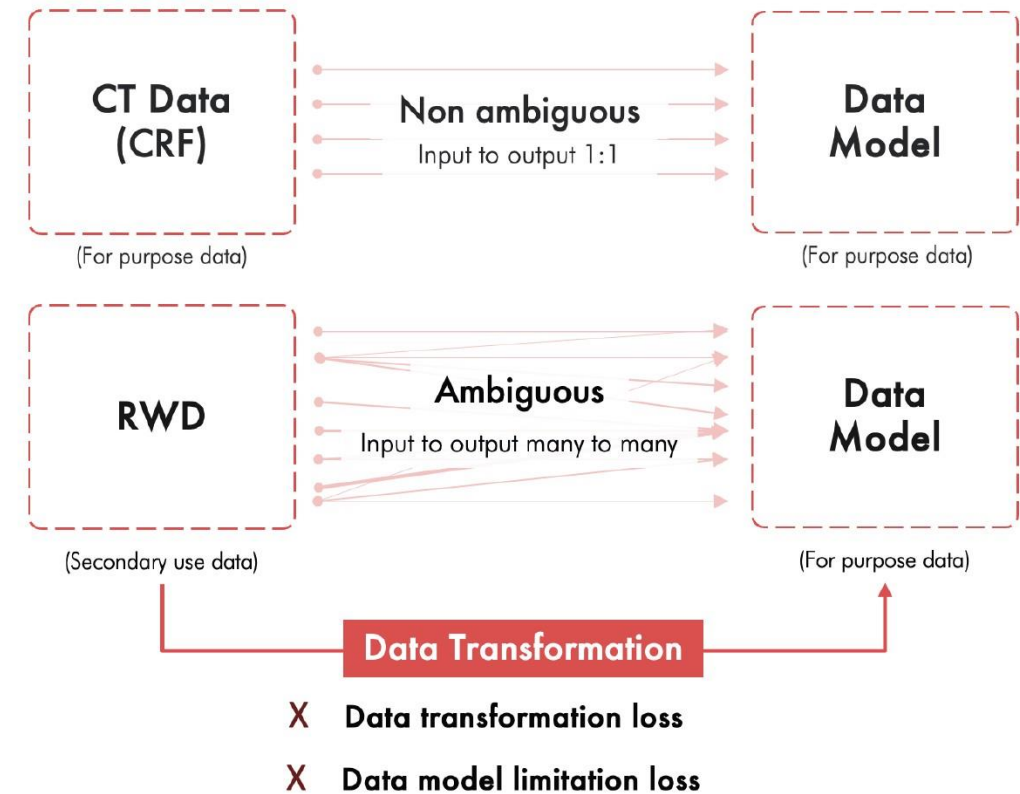
- Voidaan arvioida hoitoja ja interventioita, joista ei ole RCT-näyttöä.
- Mahdollistaa harvinaisten potilasryhmien ja alaryhmien tarkastelun.
- Pidempi seuranta-aika ja suuremmat populaatiot kuin useimmissa RCT:issä.
- Kustannustehokasta: hyödynnetään jo kerättyjä rekisteri- ja EHR-aineistoja.
- Tukee kliinistä päätöksentekoa ja sääntelyä tilanteissa, joissa RCT-dataa on vähän.
- Auttaa arvioimaan RCT-tulosten yleistettävyyttä reaalikäytäntöön

TARGET TRIAL -AJATTELU

- RCT antaa valmiin target trial -määritelmän (interventio + kontrolli)
- RWD:ssä yritetään:
 - toistaa samat sisäänotto-/poissulkukriteerit,
 - sama standardihoito,
 - sama seuranta ja päätetapahtumat
- Datan laatu ratkaisee, miten lähelle päästään

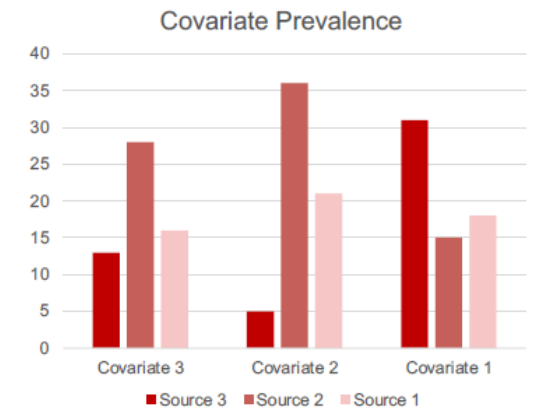
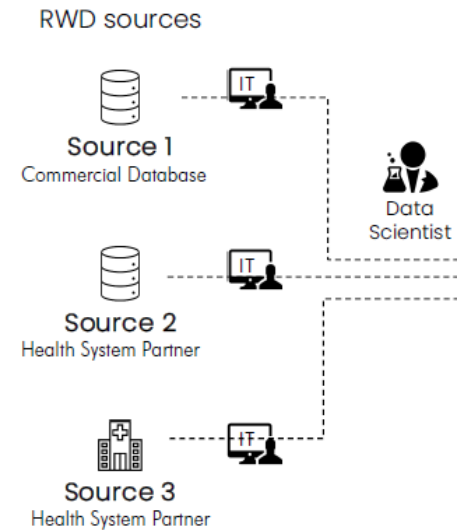
MIKSI DATAN LAATU ON ERITYISEN KRIITTINEN RCT– RWD-VERTAILUSSA?

- RCT-data: tutkimukseen suunniteltu, protokollaohjattu
- RWD-data: hoidon sivutuote
 - → puuttuva data, mittausvirhe, epäsäännöllinen seuranta
- Regulaatio:
 - Kysymykset usein ensimmäisenä datan laadusta, validoinnista ja standardoinnista



TÄRKEÄT LAADUN ULOTTUVUUDET

- Seurannan täydellisyys:
 - Kuolema, sairaalahoidot, lopputapahtumat – kuinka täydellisesti havaitaan?
- Linkitys:
 - Voiko yhdistää eri rekistereitä (esim. hoito + kuolema + diagnoosit) luotettavasti?
- Standardoitu datamalli:
 - Esim. OMOP – helpottaa monen maan / lähteen yhdistelyä
- Dokumentaatio:
 - Keruuprosessi, muutokset koodauksessa, laadunvalvonta
- Maantieteellinen ja ajallinen kattavuus:
 - Vastaavatko maat ja kalenterijaksot RCT:tä ja kohdemarkkinoita?



METODOLOGISET AVAINKOHDAT RCT–RWD-VERTAILUSSA

- Target trial -määrittely etukäteen – ei post hoc -analyysia
- "New user, active comparator" -asetelma (immortal time bias -esto)
- Confoundingin hallinta:
 - Propensity score (matching / weighting / stratification)
 - Kovariaattien tasapainon arviointi (SMD)
 - Seurannan ja sensuroinnin yhdenmukaistaminen RCT:n kanssa
- Missing data ja mittausvirheet:
 - käsittelymenetelmät pitää suunnitella jo datan hankintavaiheessa

RCT & RWD:N KÄYTTÖ LÄÄKEKEHITYKSESSÄ

- RCT: sisäinen validiteetti, satunnaistus, standardoitu datankeruu
- RWD: edustaa todellista hoitoa ja potilaita, mutta datan laatu ja rakenne vaihtelee
- Uusi lääke ei vielä markkinoilla:
 - RCT-interventiohaara: uusi lääke
 - RCT-kontrolli: standard of care tutkimusolosuhteissa
 - RWD-haara: sama standard of care reaali maailmassa
- Kysymys: kuinka hyvin RWD-kontrolli voi “jäljittää” RCT-kontrollin?

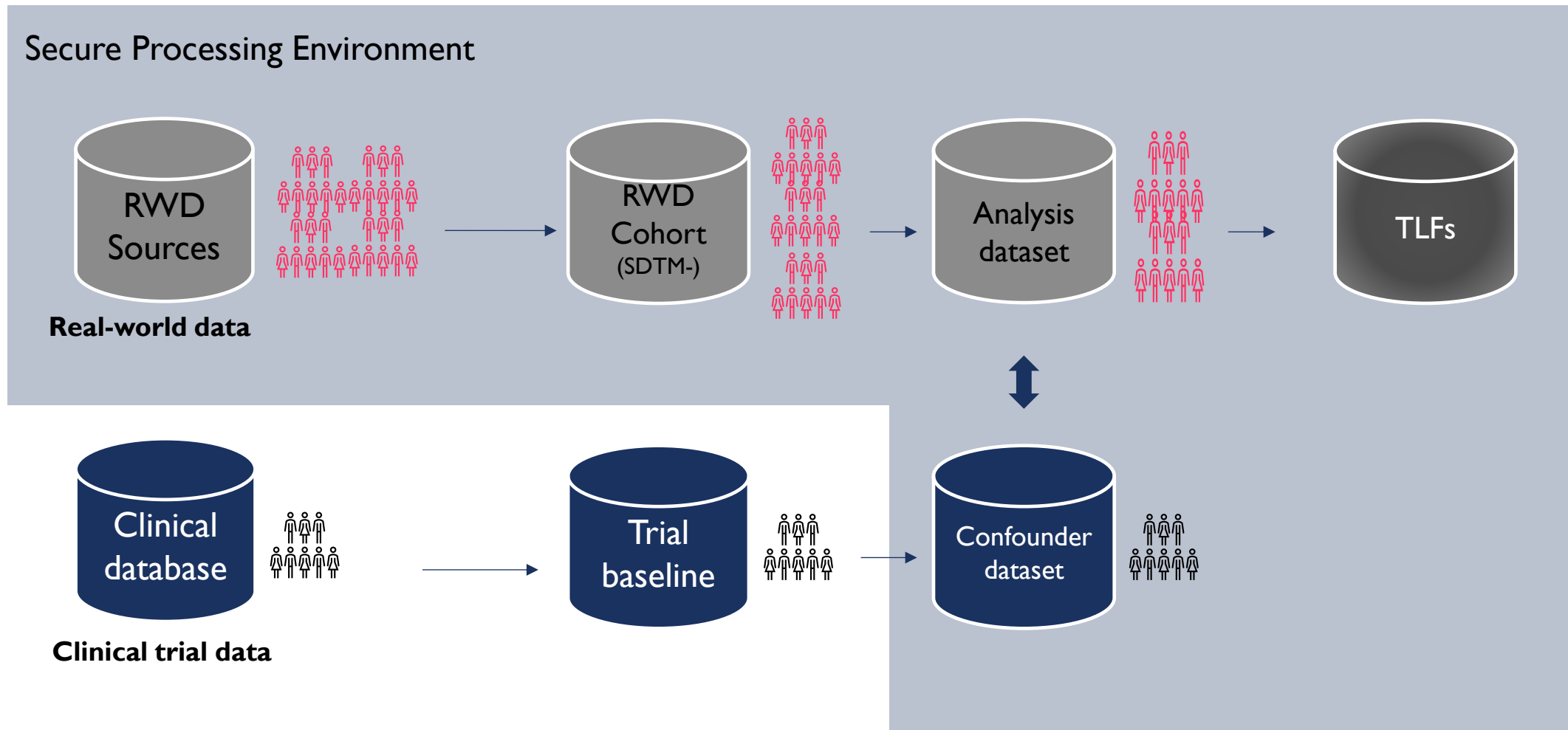
TARGET TRIAL ESIMERKKEJÄ

I. Target trial		Kidney disease (phase III)	Kidney disease (phase III)	Atrial fibrillation (phase II)
2. RWD-populaation valinta	Sovitetaan inklusio-/ekskluusiokriteerit	HUS EHR (N=23,523)	HUS EHR (N=2,646)	Auria EHR + rekisterit (N=3,327)
3. Datan harmonisointi	Kovariaattien yhdenmukaistaminen	SNOMED, ATC, ICD-10	VAE-enkoodaus	36 kovariaattia
4. Vertailukelpoisuuden arviointi	Klusterointi / matching	VAE + HDBSCAN	Latenttimuuttujamalli	PS-matching/weighting

TARGET TRIAL ESIMERKKEJÄ

1. Target trial		Kidney disease (phase III)	Kidney disease (phase III)	Atrial fibrillation (phase II)
2. RWD-populaation valinta	Sovitetaan inklusio-/ekskluusiokriteerit	HUS EHR (N=23,523)	HUS EHR (N=2,646)	Auria EHR + rekisterit (N=3,327)
3. Datan harmonisointi	Kovariaattien yhdenmukaistaminen	SNOMED, ATC, ICD-10	VAE-enkoodaus	36 kovariaattia
4. Vertailukelpoisuuden arviointi	Klusterointi / matching	VAE + HDBSCAN	Latenttimuuttujamalli	PS-matching/weighting

KÄYTÄNNÖN TOTEUTUS KONTROLLIHAAROJEN VERTAILUSSA

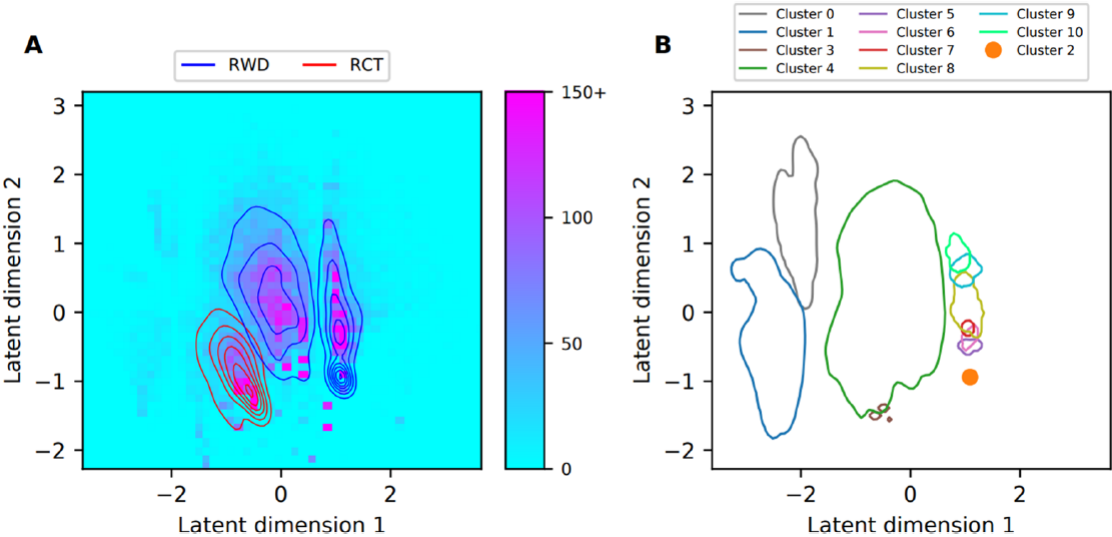


TARGET TRIAL ESIMERKKEJÄ

1. Target trial		Kidney disease (phase III)	Kidney disease (phase III)	Atrial fibrillation (phase II)
2. RWD-populaation valinta	Sovitetaan inklusio-/ekskluusiokriteerit	HUS EHR (N=23,523)	HUS EHR (N=2,646)	Auria + rekisterit (N=3,327)
3. Datan harmonisointi	Kovariaattien yhdenmukaistaminen	SNOMED, ATC, ICD-10	VAE-enkoodaus	36 kovariaattia
4. Vertailukelpoisuuden arviointi	Klusterointi / matching	VAE + HDBSCAN	Latenttimuuttujamalli	PS-matching/weighting

RCT VS RWD

	RCT	RWD
Keräystapa	CRF, protokollaohjattu	EHR, hoidon sivutuote
Pre-index -aika	20,3 v (mediaani)	5,2 v (mediaani)
Diagnoosit/vuosi	0,3	3,3
Infektiolääkkeet (J)	8,6%	66,0%

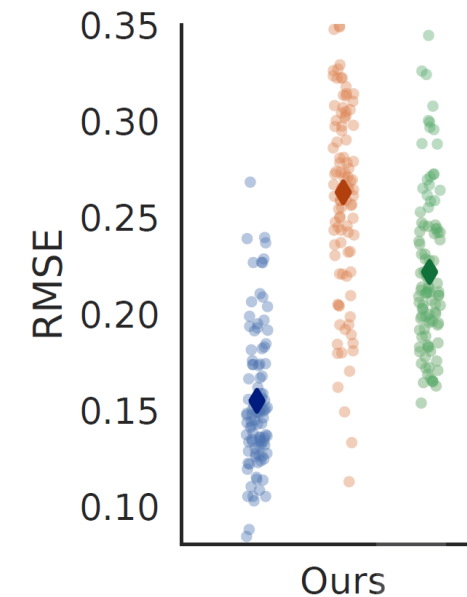
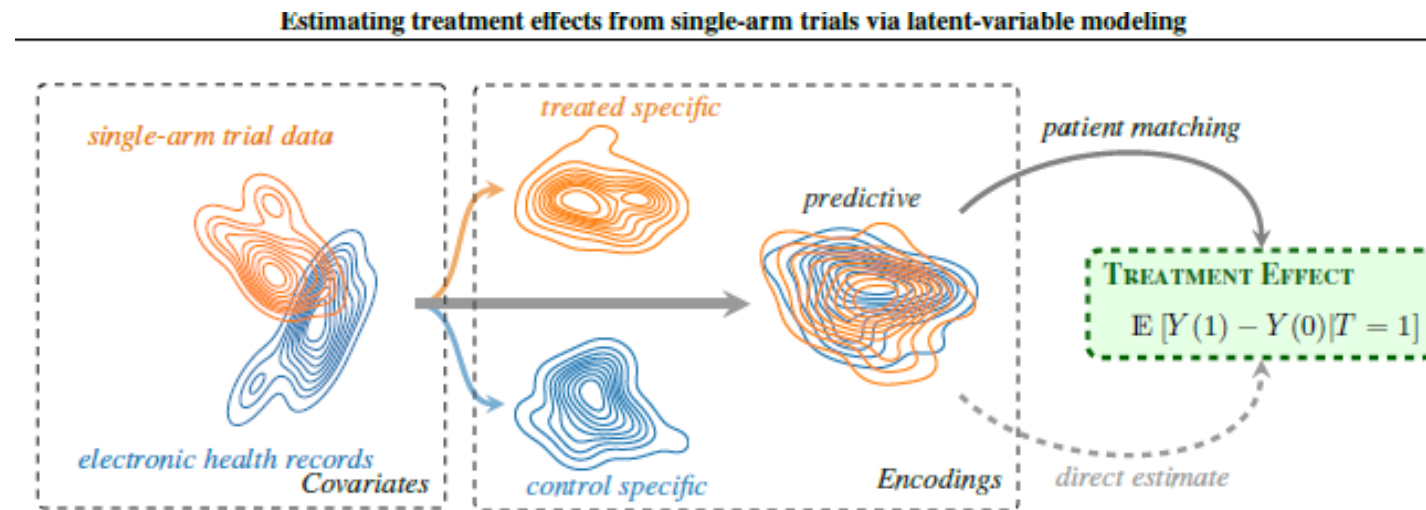


A comparative study of clinical trial and real-world data in patients with diabetic kidney disease

Samu Kurki^{1,4}, Viivi Halla-aho^{2,4}, Manuel Haussmann³, Harri Lähdesmäki³, Jussi V. Leinonen² & Miika Koskinen²

Scientific Reports | (2024) 14:1731

RCT JA RWD EROAVAT → MITEN ARVIOIDA HOITOVAIKUTUS?



Estimating treatment effects from single-arm trials via latent-variable modeling

Manuel Häußmann¹

Tran Minh Son Le¹

Viivi Halla-aho³

Samu Kurki⁴

Jussi V. Leinonen⁴

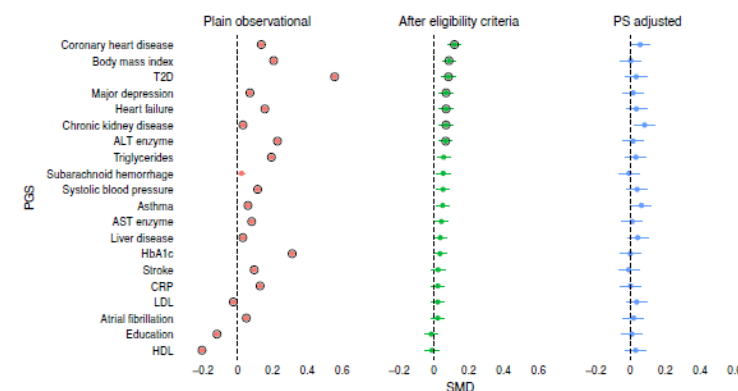
Milka Koskinen³

Samuel Kaski^{1,2}

Harri Lähdesmäki¹

GENEETTISEN DATAN INTEGROINTI TARGET TRIAL EMULATIONIIN

- **Polygenic score (PGS) mittarina:** PGS emulointiprosessin laadun arviointiin (tasapaino → hyvä confounding-kontrolli)
- **PGS-rikastaminen:** PGS:n ennustearvo emulaatiokohortissa ennen RCT-suunnittelua
 - **Prognostic enrichment (ennustava rikastaminen)**
 - Käytetään polygeenisiiä riskipisteitä (PGS) valitsemaan tutkimukseen henkilöitä, joilla on suurempi todennäköisyys saada tapahtuma (esim. sydäninfarkti).
 - Kun kaikki osallistujat ovat lähtötilanteessa korkeammassa riskissä, tapahtumia kertyy nopeammin ja pienemmällä osallistujamäärällä tai lyhyemmällä seurannalla ⇒ koe voidaan tehdä pienempänä ja halvempaan.
 - **Predictive enrichment (hoitovasteen rikastaminen)**
 - Käytetään PGS:ää valitsemaan henkilöitä, joilla on suurempi todennäköisyys hyötyä hoidosta (tai pienempi riski haitasta).



nature genetics

Article

<https://doi.org/10.1038/s41588-025-0229-8>

Incorporating genetic data improves target trial emulations and informs the use of polygenic scores in randomized controlled trial design

Received: 16 October 2024

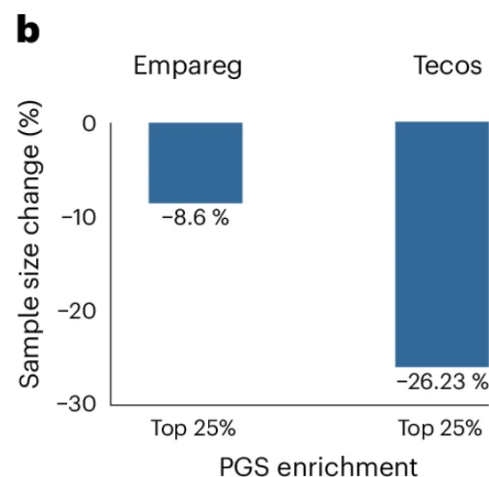
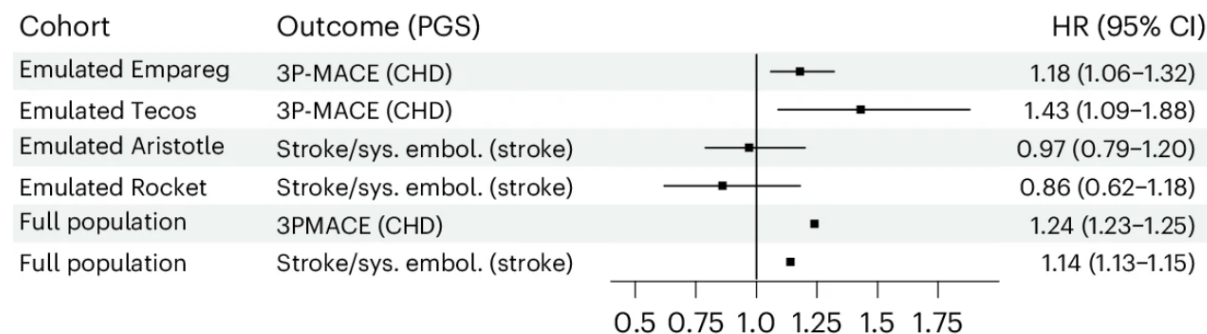
Accepted: 13 May 2025

Published online: 18 June 2025

Jakob German^{1,2}, Zhiyu Yang¹, Sarah Urbut^{1,4,5}, Pekka Vartiainen^{1,6}, FinnGen⁷, Pradeep Natarajan^{8,9,10}, Elisabetta Pattorno^{11,12}, Zoltan Kutalik^{13,14}, Anthony Philippakis^{15,16} & Andrea Ganna^{17,18,19}

YHTEENVETO

- Datan laatu ratkaisee
- Eri RWD-lähteet sopivat eri kysymyksiin ja tutkimusasetelmiin
- **Checklist:**
 - Onko populaatio rajattavissa RCT-kriteerein?
 - Onko hoidon aloitus/lopetus tunnistettavissa?
 - Ovatko päätetapahtumat vertailukelpoisia?
 - Onko seuranta riittävän täydellinen?
 - Onko datan laatu dokumentoitu?



3P-MACE Components:
Cardiovascular death
Non-fatal Myocardial Infarction (MI)
Non-fatal Stroke



KIITOS