

Työpaperi 11/2022

# Valtava-hanke – Diabetesrekisterin pääraportti

Saara Metso, Aapo Tahkola, Sini Vanhamäki, Tuuli Kauppala, Henri Salo,  
Tiina Laatikainen, Jonna Salonen, Jere Veltheim, Jaason Haapakoski

Tässä raportissa kuvataan STM:n Toivo-ohjelman Valtava-hankkeen projektin "Diabetesrekisterin tietoallaspilotti" toteutusta ja tuloksia. Projekti toteutettiin vuosina 2020–2021 yhteistyössä Valtava-hankkeen Kelan projektin "Potilastiedon arkisto tietoaaltaaseen (PTA)" sekä THL:n laaturekisterien kehittämisen projektin kanssa.

**Sote-uudistus**

# Sisällys

1 Projektin tavoitteet	4
Kirjoittajat	4
Tausta	4
Tavoite	5
Diabetesrekisterin ydinindikaattorit ja niiden valintaperusteet	5
2 Aineisto ja menetelmät	11
Populaation määrittely ja ydinindikaattorit	11
Populaation muodostaminen ja validointi	11
Laboratoriotulosten validointi	12
Alaryhmäanalyysit ja laatuindikaattorit	12
Käytettyihin tietolähteisiin kertynyt tieto ajan funktiona	14
Tietolähteiden käyttötarkoitus	18
Analyysitietokanta	19
Potilaiden identifiointi	19
Laaterekisterin minimitietosisällön mukaiset tietopoiminnot PTA-tietokannasta	21
Väestötiedot	29
ER-kaavio	29
Nimeämiskäytänteistä	30
Validoinnin tukiaineistot	32
Hilmo-tiedot	32
Fimlab-pilottiaineisto	34
Siun Sote -pilottiaineisto	35
Terveystalon työterveysasiakkaat	35
Julkisten rekisterinpitäjien kaikki PTA-laboratoriotulokset	36
Tilastolliset menetelmät	37
3 Tulokset	38
Potilaiden tunnistaminen	38
Potilaspopulaation muodostaminen ja kattavuuden arviointi	38
Alueellinen prevalenssi	41
Esimerkkejä diabeteslaaterekisteriin tuotettavista diabeteksen prevalenssia ja diabeetikoiden lukumäärää kuvaavista raporteista	43
Diabeteslääkemääräyksiin ja diagnooseihin perustuvan diabeteksen prevalenssin vertailu hyvinvointialueittain	45
Lääkitys	47
Lääkinnälliset alaryhmät	47
Muu lääkitys	49
Hoitomyöntyvyys	50
Sairastavuus ja komplikaatioiden esiintyvyys	50
Sydän- ja verisuonisairaudet	50
Retinopatia eli verkkokalvosairaus	52
Diabeteksen jalkakomplikaatiot	52
Munuaisten vajaatoiminta	53

Suusairaudet	54
Lihavuus	54
Kuolleisuus	54
Laboratoriokokeet	57
Laboratoriokokeiden tulosten kattavuudet	57
Heikon laboratoriokoe-kattavuuden syitä	65
Esimerkkejä kansallisesti skaalatusta diabetesrekisteristä saatavista hoitotasapainoa kuvaavista laatuindikaattoreista	69
4 Havainnot ja johtopäätökset	76
Potilaiden tunnistaminen	76
Toimenpiteet	78
Laboratoriotulokset	79
Tietopyyntöön lisättävät laboratoriotutkimukset	81
Fysiologiset mittaukset	82
Väestötiedot	82
Lääkitys	82
Johtopäätökset	83
5 Jatkosuunnitelma ja suositukset	84
6 Teknologialiite	86
Tausta	86
Potilastiedon arkisto PTA	86
Asiakirjan rakenne	86
CDA R2 hoitoasiakirjan rakenne	87
Asiakirjat ja niiden versiot	88
Palvelutapahtumat	89
Merkinnät	90
Entryt	90
Diagnoosit	90
Toimenpiteet	91
Laboratoriotulokset	91
Fysiologiset mittaukset	91
Analyysitietokanta	92
Nimeämiskäytänteistä	92
Diagnoosit	93
Potilaiden identifiointi	95
Lääkitys	99
Sairastavuus (comorbidity)	102
Kuolleisuus	104
Laboratoriotulokset	105
Fysiologiset mittaukset	108
Tukiaineistot	109
Julkisten rekisterinpitäjien PTA-laboratorioanalyysit	109

# 1 Projektin tavoitteet

## Kirjoittajat

Saara Metso (ylilääkäri, PSHP), Aapo Tahkola (kehittäjälääkäri, Jyväskylä), Sini Vanhamäki (tilastotutkija, THL), Tuuli Kauppala (tilastotutkija, THL), Henri Salo (tilastotutkija, THL), Tiina Laatikainen (johtaja, tutkimusprofessori, THL), Jonna Salonen (ylilääkäri, THL), Jere Veltheim (projektipäällikkö, THL), Jaason Haapakoski (erityisasiantuntija, THL)

## Tausta

Tässä raportissa kuvataan [STM:n Toivo-ohjelman Valtava-hankkeen](#) projektin "Diabetesrekisterin tietoallaspilotti" toteutusta ja tuloksia. Projekti toteutettiin vuosina 2020–2021 yhteistyössä Valtava-hankkeen Kelan projektin "Potilastiedon arkisto tietoaltaaseen (PTA)" sekä THL:n laaturekisterien kehittämisen projektin kanssa.

Vuonna 2019 voimaan tulleen sosiaali- ja terveystietojen ns. [toisiolakipaketin](#) myötä tuli mahdolliseksi käyttää Kanta-tietoja toissijaisiin käyttötarkoituksiin ja tämän myötä myös laaturekistereissä. STM on laatinut lakia ja aiheeseen liittyviä strategioita mukaillen sosiaali- ja terveystietojen kokonaisarkkitehtuurin. Kanta-tietovarantojen on toivottu toimivan jatkossa merkittävänä ja jopa pääasiallisena kansallisten sosiaali- ja terveystietojen toissijaisen tiedon lähteenä, sillä niihin tallentavat tietoja kattavasti niin julkiset kuin yksityiset palveluntuottajat. Kanta-tietovarantojen soveltuvuutta THL:n kansallisten rekisterien muodostamiseksi on selvitetty vuonna 2018 TIREKA-hankkeessa, jossa aloitettua työtä on jatkettu Valtava-hankkeen kahdessa projektissa rokotustiedoilla sekä tässä raportissa kuvatuilla diabeteslaaturekisterin tiedoilla. Myös [sosiaalihuollon](#) sekä ensihoidon kansallisten tiedonkeruiden on jatkossa suunniteltu pohjautuvan Kanta-tietovarantojen tietoihin.

Diabetespotilaita hoidetaan ja seurataan hyvin laajasti terveydenhuollossa niin perusterveydenhuollossa kuin erikoissairaanhoidossa, julkisessa ja yksityisessä terveydenhuollossa. Suurinta osaa diabeteslaaturekisterin ydinindikaattoreista, kuten diagnoosit, toimenpiteet, lääkitykset ja laboratoriotutkimukset tallennetaan Kanta-tietovarantoihin rakenteisessa muodossa. Tiedot tallennetaan [Kanta-tietovarantoihin sosiaali- ja terveydenhuollon toimintayksiköiden tietojärjestelmistä](#). Tallentamisessa pitäisi käyttää kansallisesti yhtenäisesti määriteltyjä rakenteita, koodistoja ja luokituksia. Käynnit ja hoitajaksot sekä niihin liittyvät diagnoosit, toimenpiteet ja laboratoriotutkimukset tallennetaan Potilastiedon arkistoon pääasiassa asiakas- ja potilastietojärjestelmistä. Sosiaali- ja terveydenhuollon yksiköissä laboratoriotietojen tallentamiseen käytetään erillisiä laboratoriotietojärjestelmiä ja niistä tiedot kopioidaan yleensä asiakas- ja potilastietojärjestelmiin, jotta ne tallennetaan edelleen Potilastiedon arkistoon. Lääkeresepitiedot tallennetaan Reseptikeskukseen. Diabeteksen hoidon laadun mittaamisessa tarvittaisiin lisäksi muita tietoja, kuten ns. fysiologisia mittauksia (pituus, paino, verenpaine), tupakointitieto ja potilaiden itsensä raportoimia vointitietoja.

Diabetesrekisteri on yksi THL:n kehitysvaiheessa olevista yhdeksästä kansallisesta laaturekisteristä. Diabetesrekisterin toiminta käynnistyi vuonna 2018 [THL:n Terveydenhuollon kansalliset laaturekisterit -pilottihankkeessa](#). Vuosina 2018–2020 toteutetun pilottihankkeen jälkeen THL:n kansallisen laaturekisteritoiminnan kehittämistä jatketaan [kaksivuotisessa projektissa](#).

Kansalliset laaturekisterit ovat keskenään erilaisissa kehitysvaiheissa ja niiden tämän hetken tietotuotanto on toteutettu riippuen esimerkiksi kunkin rekisterin tietosisällöstä ja tiedonkeruun historiasta. Osassa laaturekistereistä tietoja kerätään tällä hetkellä pelkästään erikoissairaanhoidosta käyttäen tähän tarkoitukseen kehitettyjä tietojärjestelmiä. Osassa

laaturekisteriä tiedon tuottaminen edellyttää ylimääräistä tietojen kirjaamista. Suurimassa osassa kansallista laaturekistereistä tiedot kootaan rekisterin tietokantaan pari kertaa vuodessa tehtävillä tietopoinnoilla. Jatkossa tavoitteena on yhtenäistää laaturekisterien tietotuotantoa ja edetä jatkuvaluonteiseen tietojen toimittamiseen sosiaali- ja terveydenhuollon organisaatioista laaturekisterien tietokantoihin. Tärkeänä tavoitteena on saada kerättyä tiedot laaturekistereihin ilman, että sote-ammattilaisten tarvitsee tehdä ylimääräistä kirjaamistyötä tiedon tuottamiseksi.

## Tavoite

Projektissa pyrittiin testaamaan kansallisen diabetesrekisterin potilaspopulaation ja valittujen ydinindikaattorien poimintaa ja raportointia käyttäen hyväksi Kanta-tietoaltaan tietosisältöä. [STM:n diabetesrekisterin esiselvitystyössä](#) esitettiin Kanta-tietovarantoja diabetesrekisterin pääasialliseksi tiedonlähteeksi.

Kanta-tietoaltaasta poimittua potilaspopulaatiota ja ydinindikaattoreita verrattiin [Pirkanmaan, Keski-Suomen ja Siun soten sairaanhoitopiirien alueelta pilottivaiheessa poimittuun tietosisältöön](#), jonka tietolähteinä olivat Hilmo- ja AvoHilmo, Fimlabin laboratoriotietokanta ja Siun soten potilaskertomusjärjestelmä. Tarkoitus oli validoida kahdesta eri tietolähteestä saadut ydinindikaattorit sekä valmistella Diabeteslaaturekisterin laajenemista koko maahan Kanta-tietovarantoja hyödyntäen. Lisäksi pyrittiin tunnistamaan Potilastiedon arkiston (PTA) tiedon kattavuutta ja puutteita sekä valmisteltiin pysyvän tietotuotannon käynnistämistä.

## Diabetesrekisterin ydinindikaattorit ja niiden valintaperusteet

Valitsimme diabetesrekisterin ydinindikaattorit käyttäen apuna IOM (Institute of Medicine) kuutta terveydenhuollon laadun elementtiä. Sen jälkeen arvioimme, mitkä valituista mittareista täyttivät jo tällä hetkellä mittareiden laatuksiteerit ja mitkä vaativat jatkotoimia ennen käyttöönottoa. Näiden kahden mittareiden valinnan näkökulman asettaminen ristiin taulukossa auttaa hahmottamaan, mitkä ovat diabeteksen hoidon laadun arvioinnin ja vertaiskehittämisen kannalta mahdollisimman nopeasti käyttöön otettavat ja mahdollisimman hyvin diabeteksen hoidon tavoitteita mittaavat mittarit.

### IOM:n terveydenhuollon laadun elementtien mukaan valitut mittarit

#### Vaikuttavuus

- Toteutuvatko kansainvälisten ja kansallisten hoitosuosituksen mukaiset sairauden hoidon tavoitteet
- Diabeteksen hoidon tavoite on ehkäistä diabeteksen komplikaatioita ja niihin liittyvää ylikuolleisuutta sekä diabeteksestä johtuvaa elämänlaadun heikkenemistä.
- Valtava-projektissa vaikuttavuuden mittareiksi valittiin pilottivaiheessa määritelty ja testattu komplikaatioiden esiintyvyys (hidas mittari) ja siihen assosioituvan sokeritasapainon mittari HbA1c, kolesterolitasapainon mittari LDL ja munuaistoiminnan mittarit eGFR ja Ualbkre. Lisää tietoa laboratorionkokeista esim. [Duodecim Terveyskirjaston Laboratoriotutkimusten tulkinta -tietokannasta](#).
- Tarkoituksena on jatkossa määritellä valituille mittareille kansallisia, populaatiotason tavoitteita, esimerkiksi 25-74-vuotiaista tyypin 2 diabeetikoista < 5 %:lla HbA1c on yli 75 mmol/mol.

#### Oikeudenmukaisuus

- Toteutuuko hoito saman tasoisena potilaan asuinkunnasta riippumatta.

- Oikeudenmukaisuutta kuvaavat vaikuttavuuden mittarit esitetään erilaisissa hoidollisissa alaryhmissä mm. diabetestyyppiin, kardiovaskulaarisairausriskiin ja ikään perustuen. Oikeudenmukaisuuden toteutumista mitataan arvioimalla kuntien tai hyvinvointialueiden välisiä eroja vaikuttavuuden ja seurannan toteutumisen mittareissa, esimerkiksi kuinka monta prosenttia 25-74-vuotiaista tyyppin 2 diabeetikoista on erittäin huonossa sokeritasapainossa, eli HbA1c yli 75 mmol/mol.

### Potilaskeskeisyys

- Otetaanko hoidossa huomioon potilaan kokemus hoidosta (PROM) ja elämänlaadusta (PREM).
- Laaturekisteri-pilotissa ja Valtava-projektissa ei ollut mahdollisuutta pilotoida potilaskeskeisyyden mittareita, koska PROM- ja PREM-mittaustietoja ei ole vielä kansallisesti ja laaja-alaisesti saatavilla.
- Esimerkkejä laaturekisteri-pilotissa pohdituista mittareista:
  - PROM: geneeriset (esim. 15D, EQ-5D) ja sairausspesifit (esim. PAID) elämänlaadun mittarit. Geneeriset, kaikissa THL:n kumppanuusrekistereissä käytettävät mittarit valitaan THL:n yhteisprojektissa 2022 loppuun mennessä. Koska diabeetikot ovat monisairaita, heidän hoitonsa on hajautettu ja heitä on paljon, mittarin tulee soveltua hyvin erilaisille potilaille mukaan lukien monisairaapotilaat, sitä pitää pystyä hyödyntämään myös potilaan hoidossa, sen tulee mahdollistaa ensisijaisesti kansallinen vertaiskehittäminen ja toissijaisesti kansainvälinen vertailtavuus ja olla käytöltään kustannustehokas. Sairausspesifin elämänlaadun mittari vaatii myös kansallisen diabeteksen hoidon ammattilaisten, potilaiden ja terveydenhuollon yksiköiden välisen konsensuksen.
  - PREM: potilaan kokemus hoidosta, esim. NPS.

### Tehokkuus

- Hukan välttäminen ja tehokkaat prosessit diabeteksen hoidossa. Voidaan tarkastella ainakin kahdesta näkökulmasta:
  - Tekninen tehokkuus: euroa/tuotettu palvelu (esim. DRG) tai käytettyjen resurssien määrä/potilas (käynnit/potilas ja laboratorio-kokeiden määrä/potilas) suhteessa saavuttuihin tuloksiin (HbA1c, LDL).
  - Kohdentava tehokkuus: euroa/QALY
- Laaturekisteri-pilotissa ja Valtava-projektissa tarkasteltiin diabeteshoidon prosessia seurannan kattavuutta kuvaavilla indikaattoreilla (kuinka moni potilaista on käynyt HbA1C, LDL, eGFR tai ualbkre-laboratoriokokeissa viimeisten 2 vuoden aikana)
- Jatkossa laajennetaan seurannan kattavuuden indikaattoreita, esim. kuinka moni potilaista on käynyt silmänpohjakuivissa viimeisen 2 vuoden aikana, sekä liitetään tarkasteluun resurssien käyttö.

### Oikea-aikaisuus

- Oikea-aikaisen diagnostiikan ja hoidon toteutuminen diabeetikoilla
- Laaturekisteri-pilotissa ja Valtava-projektissa ei vielä ole ehditty pilotoida oikea-aikaisuuden mittareita
- Esimerkki Valtava-aineistolla toteutettavissa olevasta mittarista:
  - Kuinka monella potilaista, joilla on erittäin korkea sydäntapahtuman riski (sydän- ja verisuonisairaus, albuminuria tai eGFR alle 60), on käytössä ACE-estäjä-

tai statiini- eli kolesterolilääkitys, LDL-kolesteroli <1,4mmol/l ja HbA1c <53mmol/l.

### Turvallisuus

- Turvallisen hoidon toteutuminen.
- Esimerkkejä turvallisuuden mittareista, jotka yleensä liittyvät lääkehoidon toteutuksen ongelmiin tai lääkkeiden sivuvaikutuksiin
  - Hypo- ja hyperglykemiat ja ketoasidoosit PTA:sta diagnoosien perustella
  - Hypoglykemiaa aiheuttavien lääkkeiden käyttö hyvin tiukassa sokeritasapainossa olevilla ja iäkkäillä tyypin 2 diabeetikoilla
  - Diabetekseen liittyvät käynnit ensihoidon Erica-järjestelmästä.

### Toiminnan prosessi- ja vaikuttavuusmittareiden laadun arviointi

1. **tärkeys** (*importance*): Onko mittarin arvioimaa asiaa tärkeää parantaa ja mitata hoidon laadun näkökulmasta (korkeat kustannukset, iso potilasvolyymi)?
2. **merkittävyys** eli relevanssi (*relevance*): Onko mittari olennainen sairauden hoidon tavoitteiden (elämänlaatu, komplikaatioiden eston) kannalta? Löytyykö kansainvälistä tai kansallista vertaiskehittäjää, jolla sama mittari? Onko tulokseen mahdollista vaikuttaa toiminnalla? Onko tieteellistä näyttöä siitä, että mittari assosioituu vahvasti komplikaatioihin tai elämänlaatuun?
3. **uskottavuus** / validiteetti (*validity*): Mittaako mittari sitä, mitä sen on tarkoitus mitata, **kattavasti** ja tehokkaasti. Esim. HbA1c:n paraneminen mittaa omahoidon ohjauksen onnistumista paremmin kuin hoidonohjauksen kirjauksien seuraaminen. Onko mittauskattavuus arvioitavissa ja riittävä, jotta ei synny valikoitumisharhaa? Suositetaan mittareita, joita hyödynnetään joka tapauksessa potilaan hoitoprosessissa, jolloin mittauskattavuus on riittävä.
4. **luotettavuus** / reliabiliteetti (*reliability*): Jos mittari on täysin reliabeeli, siihen eivät vaikuta satunnaisvirheet eivätkä muutokset esimerkiksi asiakasrakenteessa. Kuntakoodin perusteella poimitut laboratoriotulokset on tehty validoiduissa laboratorioissa, jolloin laboratoriotuloksiin liittyy vain vähän preanalyttisiä ja analyttisiä virhelähteitä. Luotettavuutta voidaan hallita myös hoidollisilla alaryhmillä ja vakioinnilla sekä välttämällä liian pienien potilasryhmien vertailua keskenään (sattuman vaikutus suuri). Nyrkkisääntö on, että alle 200 potilaan joukossa satunnaisvaihtelun riski on suuri.
5. **käyttökelpoisuus** (*feasibility*): Mittarin kustannustehokkuus eli kuinka paljon nähdään mittarin muodostamisen vaivaa sen tuoman tiedon hyötyyn nähden. Tämä osoittautui keskeiseksi mittareiden karsiutumiseen johtavaksi mittariksi. Valitaan mittareita, jotka syntyvät sivutuotteena ydintoiminnan ohessa eli niitä käytetään aktiivisesti ja yleisesti potilashoidossa. Lisäksi tiedon siirtyminen kansallisen laadun seurannan tarpeisiin tulee olla kustannustehokasta ja automatisoitua. Vältetään kaksoiskirjaamista pelkästään hoidon laadun seuraamisen vuoksi. Käytetään geneerisiä ja koko maahan skaalautuvia mittareita mahdollisimman paljon (esim geneerinen elämänlaatumittari, ICD-10-diagnosikoodit ja hoitosuosituksen mukaiset hoidon seurannassa käytettävät laboratoriotulokset)

## 6. TAULUKKO. Indikaattoreiden kriteerit

- X-merkki kuvaa indikaattoreiden kriteerien täyttymistä. Toistaiseksi on keskitytty kehittämään vaikuttavuuden mittareita.
- Laaturekisteritoiminnan alkuvaiheessa on syytä priorisoida niitä indikaattoreita, joissa kaikki kriteerit täyttyvät. Näin toimittiin Terveystieteiden tutkimuskeskuksen laaturekisterit -pilotissa.
  - Joidenkin laadun ulottuvuuksien kohdalla vaaditaan vielä paljon kehittämissuorityötä, ennen kuin kriteerit täyttävät indikaattorit voidaan ottaa käyttöön.

IOM ulottuvuus	Tärkeys	Merkittevyys	Uskottavuus	Luo-tetta-vuus	Käyttö-kelpoi-suus	Tietolähde	Sidosryhmä, kommentit
<b>Vaikuttavuus</b>							
<u>Kuolleisuus ja komplikaatiot</u>							
Kuolleisuus	x	x	x	x	x	Väestötietorekisteri	Insuliinipuutosdiabetes ja tyypin 2 diabetes
Kardiovaskulaarisairaudet	x	x	x	x	x	PTA tai Hilmo/AvoHilmo	Iskeemisten sydänsairauksien rekisteri Dyslipidemia-KH
Retinopatia	x	x	x	x	x	PTA tai Hilmo/AvoHilmo	Diabeettinen Retinopatia-KH ACES-RWM
Diabeettinen munuaistauti	x	x	x	x	x	PTA tai Hilmo/AvoHilmo ja labrat	Munuaistautirekisteri Diabeteksen munuaistauti-KH
Jalkahaavat	x	x	x	x	x	PTA tai Hilmo/AvoHilmo	Diabeettinen jalkahaava
Ketoasidoosi	x	x	x	x	x	PTA tai Hilmo/AvoHilmo	Insuliinipuutosdiabetes
<u>Komplikaatioihin assosioituvat muuttujat</u>							
Suusairaudet	x	x				PTA tai Hilmo/AvoHilmo	Suusairauksien rekisteri
HbA1c	x	x	x	x	x	PTA tai Fimlab/Siun sote	Insuliinipuutosdiabetes ja tyypin 2 diabetes KH
LDL	x	x	x	x	x	PTA tai Fimlab/Siun sote	Insuliinipuutosdiabetes ja tyypin 2 diabetes KH ja Dyslipidemia KH
U-albkre	x	x	x	x	x	PTA tai Fimlab/Siun sote	Insuliinipuutosdiabetes ja tyypin 2 diabetes ja Diabeettinen munuaistauti KH



IOM ulottuvuus	Tärkeys	Merkittevyys	Uskottavuus	Luo-tetta-vuus	Käyttö-kelpoi-suus	Tietolähde	Sidosryhmä, kommentit
eGFR	x	x	x	x	x	PTA tai Fimlab/Siun sote	Insuliinipuutosdiabetes ja tyypin 2 diabetes ja Diabeettinen munuaistauti KH
Elintavat: Liikunta, ravitsemus, tupakointi	x	x					
Fysiologiset mitaukset: Verenpaine, BMI	x	x					Jatkossa tavoitteena PTA
<b>Potilaskeskeys</b>	<b>Tärkeys</b>	<b>Merkittevyys</b>	<b>Uskottavuus</b>	<b>Luo-tetta-vuus</b>	<b>Käyttö-kelpoi-suus</b>		
Geneerinen elämänlaatu	x	x					THL kumppanuusrekisteriprojekti
Diabetesspesifi elämänlaatu	x	x					Diabeteksen hoitoon erikoistuneissa yksiköissä ja huonon hoitotasapainon potilailla
PREM (Kokemus hoidosta)	x	x					
<b>Oikeudenmukaisuus</b>	<b>Tärkeys</b>	<b>Merkittevyys</b>	<b>Uskottavuus</b>	<b>Luo-tetta-vuus</b>	<b>Käyttö-kelpoi-suus</b>		
Mittarikohtaiset laatuvaheet, kuntien välinen vertailu	x	x	x	x	x	PTA tai Hilmo/AvoHilmo, laboratoriojärjestelmät	
Seurannasta pudonneiden osuus	x	x	x	x	x	PTA tai Hilmo/AvoHilmo, laboratoriojärjestelmät	
<b>Oikea-aikaisuus</b>	<b>Tärkeys</b>	<b>Merkittevyys</b>	<b>Uskottavuus</b>	<b>Luo-tetta-vuus</b>	<b>Käyttö-kelpoi-suus</b>		
Komplikaatioiden ilmaantuvuus	x	x	x			PTA tai Hilmo/AvoHilmo	yhteistyö mm Iskeemisten sydänsairauksien rekisterin kanssa
Gestaatiidiabeteksen konversio DM2/ 5 v	x	x	x			PTA tai Hilmo/AvoHilmo	vaatii Kelan tiedoista lääkkeiden erityiskorvattavuuden myöntöpäivän ja ensimmäisen lääkereseptin määräyspäivän
Vastasairastuneiden komplikaatiot ja HbA1c	x	x				PTA tai Hilmo/AvoHilmo	vaatii Kelan tiedoista lääkkeiden erityiskorvattavuuden myöntöpäivän ja ensimmäisen lääkereseptin määräyspäivän

<b>IOM ulottuvuus</b>	<b>Tärkeys</b>	<b>Merkittevyys</b>	<b>Uskottavuus</b>	<b>Luo-tetta-vuus</b>	<b>Käyttö-kelpoi-suus</b>	<b>Tietolähde</b>	<b>Sidosryhmä, kommentit</b>
Oikean lääkehoidon toteutuminen riskiryhmissä	x	x					
<b>Tehokkuus</b>	<b>Tärkeys</b>	<b>Merkittevyys</b>	<b>Uskottavuus</b>	<b>Luo-tetta-vuus</b>	<b>Käyttö-kelpoi-suus</b>		
Kohdentava tehokkuus: euroa/QALY	x	x					saadaan geneerisestä elämänlaatumittarista ja estimoituna päätetapahtumista
Käytetyt resurssit vs. saavutetut hoitotulokset	x	x					
<b>Turvallisuus</b>	<b>Tärkeys</b>	<b>Merkittevyys</b>	<b>Uskottavuus</b>	<b>Luo-tetta-vuus</b>	<b>Käyttö-kelpoi-suus</b>		
Vakavat hypoglykemiat	x	x				Erica	selvityksessä
Diabeteslääkehaitat	x	x					Fimea?

## 2 Aineisto ja menetelmät

### Populaation määrittely ja ydinindikaattorit

#### Populaation muodostaminen ja validointi

Populaatio poimittiin esiintyneen diabetesdiagnoosin ICD-10: E10 - E14 perusteella sekä **potilastiedon arkistosta** (PTA, koko historiatieto 2020 asti) että **AvoHilmosta** (aikaväli 2010–2020) ja **Terveys-Hilmosta** (aikaväli 2017–2021) (Hilmo). Poiminnassa huomioitiin sekä pää- että sivudiagnoosit.

Lisäksi muodostettiin populaatio **määrättyjen diabeteslääkkeiden** (eresepti, koko historia-tieto 2020 asti) ja myönnettyjen **diabeteslääkkeiden erityiskorvausoikeuksien** (etuus, aikaväli 2010–2020) perusteella. Diabeteslääkkeiksi katsottiin ATC-luokkaan A10 kuuluvat lääkkeet. Erityiskorvausoikeuden tunnistamisessa käytettiin koodeja 103 ja 215.

Useimmista analyyseistä poistettiin henkilöt, jotka eivät löytyneet käytössä olleesta väestörekisterikopiosta (luotu 8.2.2014 elossa olevista, suomalaisen henkilötunnuksen (hetu) omaavista henkilöistä). He olivat todennäköisimmin henkilöitä, joilla väliaikainen hetu tai jotka kuolleet ennen vuotta 2014. Väestörekisterikopiota on sittemmin päivitetty säännöllisesti syntymä-, muutto- ja kuolintapahtumilla.

Lisäksi muodostettiin **kolmesta tietolähteestä (PTA, eResepti- ja etuustiedot) koottu summapopulaatio**, jota verrattiin

- (1) PTA:sta diagnoosien perusteella muodostettuun populaatioon
- (2) Hilmosta diagnoosien perusteella muodostettuun populaatioon
- (3) ereseptien perusteella muodostettuun populaatioon sekä
- (4) Kelan lääkekorvaus-etuksien perusteella muodostettuihin populaatioihin.

Populaatioita verrattiin pilottiaineistossa Siun soten, Keski-Suomen ja Pirkanmaan sairaanhoitopiirin alueella muodostettuihin populaatioihin. Tässä vertailussa on huomattava, että pilottiaineiston aikarajaus poikkeaa Valtava-hankkeessa muodostetusta populaatiosta. Pilotiaineiston määrittely:

- 1.1.2014-31.8.2021 henkilöt, jotka ovat saaneet jossain vaiheessa diabetesdiagnoosin (ICD-10: E10 - E14). Huomioidaan sekä pää- että sivudiagnoosit.
- PSHP:n ja KSSHP:n tietolähteinä ovat Hilmo ja Avohilmo. Siun Soten tietopoiminnassa on huomioitu diagnoosit myös potilastietojärjestelmän tietokannan diagnoosit-taulusta.
- Viimeinen kotikunta Keski Suomen, Pirkanmaan tai Siun soten kunnissa
- Laboratoriotutkimukset 1.1.2014 31.8.2021 Fimlabista (PSHP ja KSSHP) ja Siun sotesta potilaskertomusjärjestelmästä. Saarikan ja Wiitaunionin LDL- ja Ualbkre-tiedot on haettu CGI:n toteuttamana erillishakuna.

Yksityiskohtaisemmat tekniset määrytykset löytyvät jäljempää, kohdasta Analyysitietokanta: Potilaiden identifiointi.

## Laboratoriotulosten validointi

Osana validointia tehtiin laboratoriokoetulosten kattavuuden vertailua eri aineistojen välillä. Pilottiaineiston laboratoriokokeet oli pyydetty aikaväliltä 2015–2018 Fimlabista sekä (Fimlabiin kuulumattomista) Siun Sotesta ja Saarikasta.

## Alaryhmäanalyysit ja laatuindikaattorit

### Hoidollisten alaryhmien ja laatuindikaattorien muodostamiseksi tarvittavat muuttujat

- Ikä
- Kotikunta
- Kuolinaika
- Rekisterinpitäjä ja käytetty sairauskertomusjärjestelmä
- Diagnoosinumero (pää ja sivudiagnoosit)
- Toimenpiteet (THL Toimenpideluokitus)
- Lääkitys (ATC-koodit)
  - A10 Diabetes
  - C02-C09 Verenpaine
    - Erikseen lisäksi C09 Reniini-angiotensiinijärjestelmään vaikuttavat lääkkeet
  - B01A Veren hyytymiseen vaikuttava
  - C10 Dyslipidemia
- Laboratoriotiedot
  - Pt-GFRe, P -Krea, B -HbA1c, fP-Kol-LDL, P-Kol-LDL, U -AlbKre
- Fysiologiset mittaukset, kuten pituus, paino, verenpaine

### Projektissa toteutetut hoidolliset alaryhmät

Populaation diabeetikot jaettiin analyyseja varten **ns. hoidollisiin alaryhmiin**

- **Diabetestyyppiin perustuvat alaryhmät** (ICD-10, viimeisin diagnoosi)
  - Tyyppi 1 = E10.\*
  - Tyyppi 2 ja muu = E11.\* tai E12.\*- E14.\*
- **Ikäryhmät:** alle 25 v, 25–74 v sekä 75 v. ja yli
- **Lääkehoitoon perustuva alaryhmät** (viimeisen 12 kk aikana toteutuneet lääkeostot)
  - ATC-koodeihin A10 (diabeteslääkkeet) perustuva luokittelu:
    - i. Tabletti- tai GLP1-hoitoinen diabetes: A10B, ei A10A
    - ii. Yhdistelmähoitoinen diabetes: A10B ja A10A Alaryhmät tabletti/GLP-1 ja perusinsuliini: A10B ja A10AC tabletti/GLP-1 ja monipistoshoido: A10B ja A10AB ja A10AC tabletti/GLP-1 ja ateriainsuliini: A10B ja A10AB tabletti/GLP-1 ja yhdistelmäinsuliini: A10B ja A10AD

- iii. Insuliinipuutosdiabetes: A10A, ei A10B Alaryhmät vain perusinsuliini: A10AC vain ateriainsuliini A10AB Perus- ja ateriainsuliini A10AB JA A10AC vain sekoiteinsuliini: A10AD
- iv. Lääkkeetön hoito/elämäntapahoito: ICD-koodin mukaan diabetes, mutta ei ATC A10 mukaisia lääkkeitä.

Lisäsairastavuuteen/riskiluokitukseen perustuvia alaryhmä-analyyseja ei ehditty toteuttaa. Populaatiota ja laatuindikaattoreita tarkasteltiin lisäksi **hyvinvointialueen ja kotikunnan sekä rekisterinpitäjän käyttämän sairauskertomusjärjestelmän perusteella** tehdyn jaot-  
telun mukaan.

### Projektissa toteutetut laatuindikaattorit

- Aineistosta tehtiin joitakin raportteja laatuindikaattoreista esimerkiksi siitä, millaisia rap-  
ortteja jatkossa voidaan kansallisesti tuottaa.
  - **Kuolleisuus**
  - **Lisäsairastavuus:**
    - **Sydän ja verisuonitaudit**
      - Pää- tai sivudiagnooseina jokin tai jotkin seuraavista (potilas lasketaan ryhmään kuuluvaksi, mikäli dg löytyy milloin ta-  
hansa seuranta-ajan loppuun mennessä)
      - Sepelvaltimotauti tai sydämen vajaatoiminta: ICD-10: I20-25, I46, I50, tai FNA, FNC, FNE, FN2 alkuiset toimenpiteet.
      - Aivoverenkierron sairaudet: ICD-10: I63 - I66, (paitsi I63.6) ja PFA14, PA2VT ja PAN14 toimenpiteet
      - Jalkojen verenkierron sairaudet: ICD-10: I70.2, I79.2, E10.5, E11.5, E12.5, E13.5, E14.5 tai valtimoiden avaamistoimenpiteet PE1AT, PF1AT (PTA/Hilmo 50 %), PEF23, PE1YT, PD3YT, PF1BT, PE1BT, PEN23, PD3AT, PEH57, PEF25, PEF24, PEH56, PDQ20, PEN24, PG2YT, PFH66.
    - **Retinopatia**
      - Pää- tai sivudiagnooseina jokin tai jotkin seuraavista H36 (pois lukien H36.8), H28.0, H43.1, H45.0, H54, H40.5, H42.0, E10.3, E11.3, E12.3, E13.3, tai E14.3 tai CKC\* tai CKD\* alkuiset toimen-  
piteet
    - **Diabeteksen jalkakomplikaatiot**
      - Nilkan, jalkaterän tai alaraajan haavan revisiotoimenpiteet NHW\*, NHK99, NHS99, NHS20, QDB05, QDG20, QDG30, QDG99
      - Nilkan yläpuolinen amputaatio, toimenpiteet NFQ10, NFQ20, NGQ10, NGQ20, NHQ10, NHQ20, NFQ48 tai NGQ48
      - Nilkan alapuolinen amputaatio, toimenpiteet NHQ30, NHQ40, NHQ60.
    - **Munuaisten vajaatoiminta**
      - Luokittelu eGRF-arvon perusteella
      - Poikkeuksena ESRD (loppuvaiheen munuaisten vajaatoiminta)

- Jokin tai jotkin seuraavista diagnooseista: Hemo-dialyysi (Z49.1), peritoneaaldialyysi (Z49.2), munuais-siirre (Z94.0) tai viimeinen eGFR alle 15 ml/min/1.73m<sup>2</sup>
- **Esimerkkiraportteja laboratorikoikeista** tuotettiin pääosin vain niille hyvinvointialueille, joiden diagnoosien kirjaamisen ja laboratoriotutkimusten kattavuus katsottiin riittäväksi (yli 80 %).
  - Raportointi toteutettiin pääosin hoidollisissa alaryhmissä
  - Mittauskattavuus
    - HbA1c (Kuntaliiton (KL) laboratorionimikkeistön koodi 6128)
    - LDL-kolesteroli (KL-koodi 4599)
    - P-Krea (KL-koodi 4600)
    - U-Albkre (KL-koodi 4511)
  - Potilaat luokiteltiin saavutetun tasapainon mukaan hoitotasapainoryhmiin
  - HbA1c: viimeisin HbA1c 2 vuoden aikana <53, 54–63, 64-74 ja yli 75 mmol/mol
  - LDL-kolesteroli: Viimeisin LDL tulos 2 vuoden aikana <1,4, <1,8, 1,9–2,6 ja yli 2,6 mmol/l
  - Gfr (krea arvosta CKD-EPI kaavalla laskettuna): Viimeisin eGFR tulos 2 vuoden aikana alle 30, 30–59 ja 60 ml/min/1.73 m<sup>2</sup> tai yli

Käytettyihin tietolähteisiin kertynyt tieto ajan funktiona

### Potilastiedon arkisto (PTA)

Seuraava taulukko antaa suuntaa antavan kuvan siitä, miten Kanta-PTA:han on kertynyt asiakirjoja tietojärjestelmittäin ja vuosittain:

Tietojärjestelmä	2011		2012		2013		2014		2015		2016		2017		2018		2019		2020		Yhteensä	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Abilita Hälsovård	0	0%	0	0%	0	0%	0	0%	2431	0%	3536	0%	4890	0%	5674	0%	5921	0%	5403	0%	27855	0%
Acute	0	0%	0	0%	46	0%	7020	1%	16974	1%	24687	1%	28840	1%	32797	1%	29847	1%	140211	1%	140211	1%
AssisDent 6	0	0%	0	0%	0	0%	0	0%	0	0%	79	0%	2002	0%	8846	0%	11396	0%	22323	0%	22323	0%
DynamicHealth eArkisto	0	0%	0	0%	0	0%	0	0%	37366	3%	82411	4%	128939	5%	145471	5%	132726	5%	526913	4%	526913	4%
Effica eArkisto	0	0%	801	100%	137760	89%	521151	60%	581352	45%	640191	32%	854084	34%	1029011	35%	1033417	36%	4797767	38%	4797767	38%
Epic	0	0%	0	0%	0	0%	0	0%	0	0%	4879	0%	76417	3%	272262	10%	353558	3%	353558	3%	353558	3%
eRA	0	0%	0	0%	0	0%	0	0%	10	0%	1614	0%	2318	0%	1125	0%	5067	0%	5067	0%	5067	0%
Esko Kantaosio	0	0%	0	0%	0	0%	8586	1%	54341	4%	56049	3%	89089	4%	98640	3%	99674	3%	406379	3%	406379	3%
GFS	0	0%	0	0%	4232	3%	47620	5%	46323	4%	40348	2%	48446	2%	20376	1%	5427	0%	212772	2%	212772	2%
LKANTA	0	0%	0	0%	0	0%	0	0%	0	0%	3106	0%	87481	3%	94870	3%	91728	3%	277185	2%	277185	2%
Mediatri	0	0%	0	0%	0	0%	8389	1%	16749	1%	106159	5%	163780	6%	199620	7%	182599	6%	677296	5%	677296	5%
Multilab	0	0%	0	0%	0	0%	85748	7%	220617	11%	334978	13%	366103	13%	299704	11%	1307150	10%	1307150	10%	1307150	10%
Pegasos	142	100%	165	100%	0	0%	12438	8%	52762	6%	73359	6%	55973	3%	46524	2%	65846	2%	86514	3%	393723	3%
SoftMedic	0	0%	0	0%	0	0%	0	0%	2197	0%	4569	0%	5680	0%	5711	0%	3719	0%	21876	0%	21876	0%
Uranus	0	0%	0	0%	6	0%	226864	26%	367303	29%	764714	38%	723458	29%	738135	25%	584312	21%	3404792	27%	3404792	27%
WinHIT HITArkisto	0	0%	0	0%	0	0%	0	0%	0	0%	3090	0%	10528	0%	12788	0%	10419	0%	36825	0%	36825	0%
<b>Kaikki yhteensä</b>	<b>142</b>	<b>100%</b>	<b>165</b>	<b>100%</b>	<b>801</b>	<b>100%</b>	<b>154482</b>	<b>100%</b>	<b>874823</b>	<b>100%</b>	<b>1285248</b>	<b>100%</b>	<b>2006893</b>	<b>100%</b>	<b>2535996</b>	<b>100%</b>	<b>2902870</b>	<b>100%</b>	<b>2850272</b>	<b>100%</b>	<b>12611692</b>	<b>100%</b>

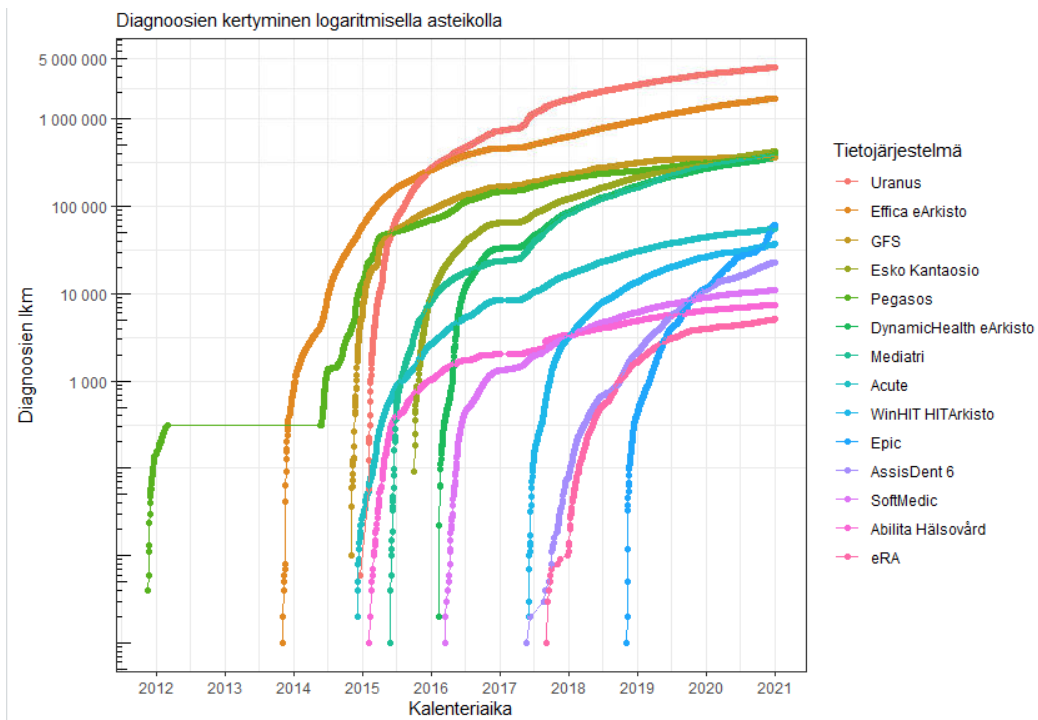
Seuraava taulukko puolestaan kuvaa vastaavalla tavalla sitä, mistä vuodesta lähtien eri tietojärjestelmät ovat tallentaneet laboratoriovastauksia tietokantaan:

Tietojärjestelmä	2014		2015		2016		2017		2018		2019		2020		Yhteensä	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Abilita Hälsovård	0	0%	1414	0%	2574	1%	3625	0%	4237	0%	4472	0%	4350	0%	20672	0%
Acute	5	0%	2649	1%	6759	1%	10199	1%	11777	1%	13515	1%	15006	1%	59910	1%
DynamicHealth eArkisto	0	0%	0	0%	4210	1%	30483	4%	42779	4%	45449	4%	41986	3%	164907	3%
Effica eArkisto	78345	100%	316137	99%	371748	79%	465519	59%	525888	48%	627347	49%	627966	48%	3012950	56%
Epic	0	0%	0	0%	0	0%	0	0%	2810	0%	43268	3%	146226	11%	192304	4%
LKANTA	0	0%	0	0%	0	0%	3106	0%	87481	8%	94870	7%	91728	7%	277185	5%
Mediatri	0	0%	0	0%	0	0%	47241	6%	83834	8%	83659	7%	82342	6%	297076	6%
Multilab	0	0%	0	0%	85748	18%	220617	28%	334978	31%	366103	29%	299704	23%	1307150	24%
SoftMedic	0	0%	0	0%	903	0%	2674	0%	2871	0%	2857	0%	1884	0%	11189	0%
<b>Kaikki yhteensä</b>	<b>78350</b>	<b>100%</b>	<b>320200</b>	<b>100%</b>	<b>471942</b>	<b>100%</b>	<b>783464</b>	<b>100%</b>	<b>1096655</b>	<b>100%</b>	<b>1281540</b>	<b>100%</b>	<b>1311192</b>	<b>100%</b>	<b>5343343</b>	<b>100%</b>

Diagnoosien kertyminen tietojärjestelmittäin näyttää seuraavalta:

Tietojärjestelmä	2011		2012		2013		2014		2015		2016		2017		2018		2019		2020		Yhteensä	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Abilita Hälsovård	0	0%	0	0%	0	0%	25	0%	1017	0%	962	0%	1265	0%	1437	0%	1449	0%	1053	0%	7183	0%
Acute	0	0%	0	0%	0	0%	0	0%	2530	0%	5573	1%	7799	1%	13643	1%	13587	1%	10470	1%	53627	1%
AssisDent 6	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	78	0%	1981	0%	8832	1%	11373	1%	22264	0%
DynamicHealth eArkisto	0	0%	0	0%	0	0%	0	0%	0	0%	33156	4%	51928	4%	86160	6%	100022	6%	90740	6%	362006	5%
Effica eArkisto	0	0%	797	100%	58418	78%	200055	37%	201667	26%	164139	26%	319407	23%	391544	25%	395583	28%	1731610	25%	1731610	25%
Epic	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	324	0%	7973	1%	45832	3%	54129	1%	54129	1%
eRA	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	10	0%	1614	0%	2317	0%	1124	0%	5065	0%
Esko Kantaosio	0	0%	0	0%	0	0%	8304	2%	51573	7%	51519	7%	86382	6%	96549	6%	97719	7%	392046	6%	392046	6%
GFS	0	0%	0	0%	4232	6%	47620	9%	46323	6%	40348	6%	48446	3%	20376	1%	5427	0%	212772	3%	212772	3%
Mediatri	0	0%	0	0%	0	0%	7941	1%	15547	2%	57865	5%	78793	6%	114951	7%	99395	7%	374492	5%	374492	5%
Pegasos	142	100%	165	100%	0	0%	12409	17%	52555	10%	73105	9%	55713	5%	46292	3%	65553	4%	84881	6%	390815	6%
SoftMedic	0	0%	0	0%	0	0%	0	0%	1294	0%	1895	0%	2809	0%	2854	0%	1835	0%	10687	0%	10687	0%
Uranus	0	0%	0	0%	6	0%	218532	41%	352937	45%	728763	63%	702370	50%	712421	46%	566603	40%	3287532	47%	3287532	47%
WinHIT HITArkisto	0	0%	0	0%	0	0%	0	0%	0	0%	3090	0%	10528	1%	12788	1%	10419	1%	36825	1%	36825	1%
Kaikki yhteensä	142	100%	165	100%	797	100%	75090	100%	538554	100%	782137	100%	1164412	100%	1401086	100%	1556216	100%	1422454	100%	6941053	100%

Seuraava kaavio havainnollistaa diagnoosien kertymistä Potilastiedon arkistoon tietojärjestelmittäin 2010-luvulla:



### Erityiskorvausoikeudet

Diabeteslääkkeiden erityiskorvausoikeuksia pyydettiin aineistoon vuodesta 2010 lähtien. Näin ollen ensimmäiset myöntämivuodet alkavat vuodesta 2010. Huomattavaa kuitenkin on, että vuoden 2016 säädosmuutos etuuskorvauksissa sai aikaan merkittävän lisäyksen koodilla AE215 oikeutettujen diabetespotilaiden määrään.

Koodi	Vuosi	Lkm
103	2010	24 329
103	2011	33 491
103	2012	26 328
103	2013	26 093
103	2014	24 758
103	2015	24 015
103	2016	23 847
103	2017	21 782
103	2018	22 532

Koodi	Vuosi	Lkm
103	2019	22 296
103	2020	15 927
215	2016	32 5393
215	2017	23 562
215	2018	22 525
215	2019	22 301
215	2020	15 927

### Reseptivuodet

eResepti otettiin käyttöön v. 2010 ja se muuttui pakolliseksi vuonna 2017. Seuraava taulukko kuvaa aineiston eReseptien kirjoitusvuosia:

Vuosi	Reseptejä
2010	991
2011	35 640
2012	474 342
2013	1 487 671
2014	1 989 706
2015	2 212 735
2016	2 329 815
2017	2 392 679
2018	1 707 054
2019	1 953 024
2020	1 748 251

Vuoden 2017 jälkeen todettava kirjoitettujen reseptien lukumäärän pieneneminen johtunee lääkemääräysten voimassaoloajan pitenemisestä kahteen vuoteen v. 2017 alusta. Ks. myös [Decreasing Number of Medication Prescriptions After e-Prescriptions Became Mandatory and Their Valid Period Was Extended: A Big Bang Policy Change in Finland in 2017.](#)



### Kuolleisuus

THL:n kopio DW:n Väestötietojärjestelmästä (VTJ) kattaa kuolemat vain 8.2.2014 lähtien. Seuraava taulukko kuvaa koko diabetesaineiston kuolinvuosien jakauman:

Vuosi	Kuolleita
2014	7 937
2015	11 736
2016	13 402
2017	14 358
2018	15 096
2019	15 305
2020	16 414
2021	3 002

Aineistoon oli kertynyt jollain kriteerillä yhteensä 8213 sellaista diabetespotilasta, jotka eivät löytyneet em. VTJ:n kopiosta. Analyysimme mukaan näistä noin puolet ovat henkilöitä, jotka olivat kuolleet ennen 8.2.2014. Muut ovat näkemyksemme mukaan väliaikaisella henkilötunnuksella kirjattuja potilaita. Poistimme analyyseistä kaikki ne henkilöt, jotka eivät löytyneet em. VTJ-kopiosta.

### TerveysHilmo (ESH)

Erikoissairaanhoidon Hilmo-tietokannasta oli poimittu tähän vertailuun aineisto, joka sisälsi vasta lähinnä 2018 alkaneita hoitajaksoja:

Vuosi	Hoitajaksoja	Potilaita
2017	17995	17167
2018	12152891	2215237
2019	12478000	2292044
2020	13852364	2229970

Aineiston analyyseistä poistettiin ne henkilöt, jotka eivät löytyneet käytettävissä olevasta Väestötietojärjestelmän kopiosta per 12.3.2021. Tarkemman analyysin perusteella näistä noin puolet oli kuollut ennen vuotta 2014.

**AvoHilmo (PTH)**

Perusterveydenhuollon hoitojärjestelmään kertyneiden tietojen kopio tässä kokonaisaineistossa jakautui hoitojakson alkamisen vuoden mukaan seuraavasti:

<b>Vuosi</b>	<b>Hoitojaksoja</b>	<b>Potilaita</b>
2010	6 264	5 700
2011	22 324	15 968
2012	12 629	7 728
2013	9 738	5 479
2014	11 188	6 505
2015	70 444	35 417
2016	217 264	56 748
2017	288 875	101 066
2018	2 931 664	837 287
2019	80 751 577	4 378 176
2020	66 978 319	4 487 841

**Tietolähteiden käyttötarkoitus**

Diabetespotilaiden tunnistamiseen käytettiin ensisijaisesti potilastiedon arkistoa (PTA), eResepti-kantaa ja lääkkeiden erityiskorvausoikeuksien etuuskantaa. Pilotissa näitä tietolähteitä vielä täydennettiin niillä potilailla, joilta löytyi DM ICD-10-diagnoosi erikoissairaanhoidon tai perusterveydenhuollon hoitoilmoitusjärjestelmistä. Tarkoituksena oli arvioida sitä, pystyvätkö kolme ensin mainittua tietolähdettä yksin kattamaan riittävästi diabetesta sairastavan väestön.

Toisena validointina käytettiin Avohilmon ICPC2-diagnoseja. Yksinomaan ICPC2-diagnoosin saaneet potilaat kuitenkin hylättiin sittemmin pääsääntöisesti analyyseistä.

Varsinaisen analyysitietokannan ydin – sairastavuus ja laboratoriotulokset – koostuu yksinomaan potilastiedon arkistosta jalostetusta aineistosta. Näin ollen esimerkiksi diabeteksen tyyppi on tiedossa vain näiltä PTA-potilailta, samoin oheissairastavuus.

Hilmon diagnoosi- ja toimenpidetietoja hyödynnettiin yksin omaan PTA-diagnoosi- ja toimenpidetietojen validoinnissa. Varsinaiset sairastavuutta ja diabeteksen komplikaatiota edustavat tunnusluvut laskettiin pelkästään PTA-diagnoosien ja toimenpiteiden pohjalta.

Reseptikeskuksesta saatavat diabetes-, verenpaine-, lipidi- ym. lääkemääräykset ja -toimitukset sen sijaan olivat käytettävissä koko aineistolle. Näin esimerkiksi lääkinnällinen alaryhmitys (insuliini-, tabletti-, yhdistelmä-lääkitys) oli tiedossa koko aineistolle.

Samoin väestötiedot (sukupuoli, syntymäaika, viimeinen kotikunta, mahdollinen kuolinpäivä) olivat käytettävissä koko aineistolle Väestötietojärjestelmän alkuvuoden 2021 tilanteen mukaisesti. Ennen 8.2.2014 kuolleiden potilaiden väestötietoja ei ollut käytettävissä. Väestötietoja käytettiin potilaan iän arvioimiseksi, alueellisten jakaumien pohjana sekä kuolleisuuden arvioinnissa.

Lisäksi pilotissa oli käytettävissä suoraan tietojen alkulähteiltä Pirkanmaan ja Keski-Suomen sairaanhoitopiirien Fimlab-laboratoriolta tilaamat HbA1c-, LDL-, kreatiniini- ja albumiini-näyteanalyysit tuloksineen, samoin Siun soten alueen vastaavat analyysit ja vielä kolmanneksi Terveystalon työterveysasiakkailta mitatut tulokset. Näitä aineistoja voitiin hyödyntää validoitaessa PTA:n laboratoriotietojen kattavuutta ja oikeellisuutta.

Seuraavassa aineisto on käsitteellisesti jaettu kahteen osaan: yhtäältä analyysitietokantaan ja toisaalta validoinnin tukiaineistoihin. Analyysitietokannaksi tässä nimitetään sitä kokonaisuutta, joka on jalostettu potilastiedon arkiston potilasasiakirjoista, reseptikeskuksen eResepti-kannasta sekä etuustiedoista. Analyysitietokannan potilaiden taulua on kuitenkin pilotissa laajennettu siten, että se sisältää tunnistetiedot myös pelkästään Hilmo-aineistosta löytyvistä potilaista tai diabetespäilyistä. Tämä ratkaisu sen vuoksi, että validoinnin tukiaineistojen käyttäminen vertailussa olisi teknisesti suoraviivaisempaa.

Kaikkien aineistojen henkilötunnisteet on kryptattu samalla algoritmilla, mikä mahdollistaa pseudonymisoidujen analyysitietokannan ja tukiaineiston vertailun myös henkiötasolla.

## Analyytitietokanta

### Potilaiden identifiointi

Potilaiden tunnistamisen pohjana oli kuusi eri tietolähdettä:

1. PTA:n ICD-10-diagnoosit (E10–E14, O24.4)
2. eResepti-kannasta lääkitysmääräykset (ATC A10)
3. Kelan erityiskorvausoikeudet (YE 103 ja AE 215)
4. Terveystalon ICD-10 (E10–E14, O24.4)
5. Avohilmon ICD-10 (E10–E14, O24.4)
6. Avohilmon ICPC2 (T89, T90, W85)

Tietokanta sisältää kaikkien näiden lähteiden mukaan tunnistetut diabetespotilaat. ICPC2 yksinään käytettynä edustaa usein vasta diabetespäilyä. Näin ollen potilaat, joilla **ainoana** diabetestietona on Avohilmon ICPC2-merkintä, on jätetty **analyyseistä pois**, ellei toisin ole mainittu.

ICD-10-diagnoosien perusteella aineisto jaettiin diabetestyyppeihin seuraavasti:

1. Tyypin I diabetes: E10\*
2. Tyypin II diabetes: E11\*
3. Raskausajan diabetes: O24.4
4. Muu diabetes: E12\*–E14\*

Useimmista analyyseistä raskausdiabetes jätettiin pois, koska se on pre- eli esidiabeteksen muoto ja sitä hoidetaan paljon ilman lääkkeitä ja erityiskorvattavuutta. Lisäksi analyyseissä yleensä yhdistettiin kohdan 4 tapaukset tyypin II diabetespopulaatioon.

PTA-aineistossa pohjana oli PTA:n fakta F\_DIAGNOOSI, jonka sarakkeista *Diagnoosin-Syy\_ICD10PK* ja *UlkoinenSyy\_ICD10PK* asianmukaisia koodilyhenteitä etsittiin säännöllisten lausekkeiden avulla (ks. [teknologia-liite](#)).

Lääkemääräysten ATC-koodin perusteella poimittiin aineistoon ne henkilöt, joille oli määrätty ATC-luokituksen A10 mukaisia lääkkeitä ajalla 2010–2020.

Aineistoon lisättiin kaikki ne henkilöt, joilla oli erityiskorvausoikeuspäätös 103 (insuliinilääkitys) tai 215 (muu kuin insuliini) ajalta 2010–2020. Huomattakoon, että kun erityiskorvausoikeus 215 otettiin käyttöön 1.1.2017, kyseisellä ajanhetkellä elossa olleet, diabeteskorvausoikeuden 103 omanneet henkilöt saivat automaattisesti myös tämän uuden koodin ([Valtioneuvoston asetus vaikeiden ja pitkäaikaisten sairauksien lääkehoidon kustannusten korvaamisesta 1149/2016 \(finlex.fi\)](https://www.finlex.fi/vaston/asetus/vaikeiden-ja-pitkaaikaisten-sairauksien-laakehoidon-kustannusten-korvaamisesta-1149/2016)).

Analyyysien pohjaksi muodostettiin taulu A\_PATIENT, joka muodostuu edellä kuvattujen neljän tietolähteen unionista. Taulun sarakkeet ovat:

Nimi	Tietotyyppi	Selitys
src	integer	Tietolähteet kakkosen potenssien $2^{**}(i-1)$ summana; i on yllä mainittu tietolähteen numero 1–6
dgid	guid	Viite siihen PTA:n ICD-10-diagnosiin (a_diagnosi), johon poiminta nojautuu; vain PTA-aineisto (lähde 1)
pid	raw(16)	Potilaan pseudotunniste
asiakirjaoid	oid	Diagnosipotilasasiakirjan tunniste; vain PTA-aineisto, viite tauluun F_YhteinenHeaderTieto
dgdate	date	Varhaisin diagnosin toteamispäivä tai kirjauspäivä kaikki tietolähteet huomioon ottaen
regdate	date	Varhaisin kirjauspäivä (rekisteröintipäivä) PTA-, lääkitys- ja etuustietolähteet huomioon ottaen
bdate	date	Potilaan syntymäaika
gender	varchar2(6)	Potilaan sukupuoli ("mies", "nainen")
dmtree	integer	Diabeteksen tyyppi (1,2,3,4); vain PTA-aineisto
rekisterinpitajaid	oid	Julkisen sektorin rekisterinpitäjän tunnus (1. diagnosi); vain PTA-aineisto, viite tauluun d_rekisterinpitaja
rekisteripitaja_organisaatioid	oid	Yksityisen rekisterinpitäjän organisaatiotunnus; vain PTA-aineisto, viite tauluun d_organisaatio
deathdate	date	Potilaan kuolinpäivä, jos tiedossa
prid	raw(16)	CDA_ID-viite lääkemääräykseen (taulu a_prescription); vain ne, joilta lääkitystiedot (lähde 2)
medtypeid	integer	Potilaan lääkityksen luokittelu (ks. lääkitys); vain ne, joilta lääkitystiedot (lähde 2); viite tauluun d_medtype
reimbursement	integer	Erytyiskorvausoikeus (1=insuliini 103, 2=muu kuin insuliini 215, 3=kummatkin; 0=ei etuustietoa);
kuntaid	varchar2(3)	Potilaan viimeisin tiedossa oleva asuinkunta (Kuntaliiton koodisto); viite tauluun d_area

Laaturekisterin minimitietosisällön mukaiset tietopoiminat PTA-tietokannasta

### Sairastavuus: diagnoosit ja toimenpiteet

#### Diagnoosien ja toimenpiteiden määrittely

Seuraava taulukko listaa diagnoosien ja toimenpiteiden mukaan luokitellut sairaudet ja toimenpiteet. Nämä sairaudet ovat diabeteksen komplikaatioita tai ne edustavat muuta oheissairastavuutta, joka tulee ottaa analyysissä huomioon esim. hoidollisena alaryhmänä.

Diagnoosi tai toimenpide	disid	Säännöllinen lauseke
<i>Diagnoosit</i>		
<b>Ketoasidoosit</b>	1	E10.1 E11.1 E12.1 E13.1 E14.1 E10.09 E11.09 E11.01
<b>Hyperglykemiat</b>	2	E10.00 E11.00 E12.0 E13.0 E14.0
<b>Infektiot</b>	3	[AB]\d{2} G0\d H60.[0123] H6[68] I J 02 J8[56] K6[567] K35 L0[012345678] M0[0123] N[13]0 N39 H60.9 K75.0 T81.4
<b>Retinopatia</b>	4	H36.[012345679] H43.1 H45.0 H54 H40.5 H42.0 E1[01234].3
<b>Diabeettinen neuropatia</b>	5	M14.[26] G59.0 G63.2 G73.0 G99.0 N48.4 E1[01234].[456]
<b>Diabeettinen jalkahaava</b>	6	L97
<b>Alaraaja-ateroskleroosi</b>	8	I7[09].2 E1[01234].5
<b>Jalkaterän lievät infektiot ja iho-ongelmat</b>	11	B35.[13] B07 L03.0 L28 L84 L85.3
<b>Jalan vakavat infektiot</b>	12	L03.1 M86.6 T81.4
<b>Jalkaterän virheasennosta ja neuropatiasta johtuvat ortopediset vaivat</b>	13	L60.0 M72.2 M20.[124] M21.[45]
<b>Aikaisempi amputaatio</b>	14	Z89.[4567]
<b>Diabetes ja munuaissairaus</b>	15	N08.3 N18 E1[01234].2
<b>Dialyysihoito</b>	16	Z49 Z94.0
<b>Iskeeminen sydänsairaus</b>	20	I2[012345] I46
<b>Sydämen vajaatoiminta</b>	21	I50
<b>Eteisvärinä</b>	22	I48
<b>Aivoverenkiertohäiriöt</b>	25	I63.[012345789] I6[456] G45
<b>Alaraaja-ateroskleroosi</b>	26	I7[09].2 E1[01234].5

Diagnoosi tai toimenpide	disid	Säännöllinen lauseke
Lihavuus	30	E66
Uniapnea	31	G47.3
Rasvamaksa tai NASH	32	K76.0 K73.80
Suun terveys: ien- ja juuri-infektiot	40	F33.[123] F34
Masennus	41	F32.[123456789] F32.3
Dementia	42	F0[0123] F1[01].73
Psykoosit, kaksi-suuntainen mielialahäiriö tai älyllinen kehitysvammaisuus	43	F20 F31 F7[0123456789]
Syömishäiriöt	44	F50
Alkoholi- tai huumeriippuvuus	45	F1[0123456789]
Kilpirauhasen sairaudet	50	E0[01234567]
Muiden umpirauhasten sairaudet	51	E16 E2[0123456789] E3[012345]
Tulehdulliset suolisairaudet	52	K5[012]
Tulehduksellinen maksa- tai sappitie-tulehdus	53	K73 K74.3 K75.4
Keliakia	54	K90.0
Tulehdukselliset nivelsairaudet	55	M0[56789] M1[01234]
Osteoporoosi	56	M8[01]
<i>Toimenpiteet</i>		
Retinopatian hoito	4	CKC CKD
Nilkan, jalkaterän ja alaraajan haavan revisiot	7	NHW NHK99 NHS99 NHS20 QDB05 QDG[23]0 QDG99
Nilkan yläpuolinen amputaatio	9	N[FGH]Q[12]0 N[FG]Q48]
Nilkan alapuolinen amputaatio	10	NHQ[346]0
Silmänpohjakuvaus	19	CL2[LS]4
Sepelvaltimoiden ohitusleikkaukset ja stenttaukset	23	FN[ACE2]

Diagnosi tai toimenpide	disid	Säännöllinen lauseke
Aivovaltimoiden toimenpiteet	24	P[AB]
Valtimokovettumataudin hoito	27	P[DEFG]
Lihavuusleikkaus	33	JDF
Suun terveys: ien- ja juuri-infektiot	34	K0[45]

Analyysitietokannan taulut *dm\_disease\_icd10* ja *dm\_disease\_proc* sisältävät kaikki ne ICD-10-diagnosikoodit ja THL Toimenpideluokituksen toimenpidekoodit, jotka vastaavat yllä olevia säännöllisiä lausekkeita. Taulujen avainkenttä *disid* on sama kuin yo. taulukon toinen sarake.

### Analyyssitaulu A\_DIAGNOOSI

Yllä lueteltuja diagnooseja vastaavat diagnosoinnit yhdessä diabetesdiagnosointien (E1[01234]|O24.4) kanssa on koottu analyysitietokannan tauluun A\_DIAGNOOSI. Sen sarakkeet ovat seuraavat:

Nimi	Tietotyyppi	Selite
dgid	guid	Diagnosille annettu yksikäsitteinen surrogaattivain (taulun primariivain)
pid	raw(16)	Potilaan pseudonymisoitu tunniste
bdate	date	Potilaan syntymäaika
gender	varchar2(6)	Potilaan sukupuoli
dgdate	date	Diagnoosin toteamispäivä tai sen puuttuessa kirjauspäivä
enddate	date	Diagnoosin lopetuspäivä, jos kyseessä diagnoosin päättäminen
syy_icd10	nvarchar2(50)	Diagnoosin syykoodi, voi olla yhdistelmädiagnosi
ulksyy_icd10	nvarchar2(50)	Diagnoosin ulkoisen syyn koodi, voi olla yhdistelmädiagnosi
toteaja_organisaatioid	oid	Toteajaorganisaation SOTE-organisaatiotunnus (viite tauluun d_organisaatio)
asiakirjaoid	oid	Potilasasiakirjan tunnus, viite mm. tauluun F_YhteinenHeaderTieto
merkintaoid	oid	Potilasasiakirjan merkinnän tunnus
entryoid	oid	Diagnoosimerkinnän tunnus; asiakirjaoid, merkintaoid ja entryoid muodostavat diagnosoinnin luonnollisen avaimen, viitteen faktatauluun f_diagnosi
alkuperainen-asiakirjaoid	oid	Ensimmäisen asiakirjaversioasiakirjatunnus (ns. setld)
pt_asiakirjaoid	oid	Viite lähimpään palvelutapahtumaan. Tämä on päätelty heuristisesti palvelutapahtumatunnuksen ja asiakirjojen aikaleimojen perusteella.
rekisterinpitajaid	oid	Julkisen rekisterinpitäjän tunnus rekisterinpitäjien rekisterissä (viite dimensioon d_rekisterinpitaja)

Nimi	Tietotyyppi	Selite
rekisteripitaja_organisaatioid	oid	Yksityisen rekisterinpitäjän SOTE-organisaatiotunnus (viite dimensioon d_organisaatio)

Itse asiassa aiemmin kuvattu taulu A\_PATIENT on osin luotu tämän taulun pohjalta, ja sen vierasavain *dgid* viittaa juuri tähän tauluun silloin, kun henkilö on tunnistettu diabetespotilaaksi PTA ICD-10-diagnoosin perusteella.

### Toimenpiteet

Vastaavalla tavalla kuin taulu A\_DIAGNOOSI on luotu taulu A\_TOIMENPIDE, johon on kirjattu kaikki ne potilaalle tehdyt toimenpiteet, joiden toimenpideluokka löytyy taulusta *dm\_disease\_proc* (ks. yllä olevat 10 toimenpideluokkaa). Sen sarakkeet ovat seuraavat:

Nimi	Tietotyyppi	Selite
procid	guid	Toimenpiteelle annettu yksikäsitteinen surrogaattivain (primaariavain)
pid	raw(16)	Potilaan psedonymisoitu tunniste
procdate	date	Toimenpiteen suorituspäivämäärä
organisaatioid	oid	Viite toimenpiteen suorittajaorganisaatioon (d_organisaatio)
noper	integer	Toimenpiteen tekijöiden lukumäärä
toimenpideluokkaid	nvarchar(5)	Toimenpideluokituksen koodi (ns. Nomesco-koodi), viite dimensiotauluun d_toimenpide
asiakirjaoid	oid	Potilasasiakirjan tunnus, viite mm. tauluun F_YhteinenHeaderTieto
merkintaoid	oid	Potilasasiakirjan merkinnän tunnus
entryoid	oid	Toimenpidemerkinän tunnus; asiakirjaoid, merkintaoid ja entryoid muodostavat toimenpiteen avaimen, viitteen faktatauluun f_toimenpide
asiakirjaluontiaika	timestamp	Toimenpidemerkinän luontiaika
pt_asiakirjaoid	oid	Viite lähimpään palvelutapahtumaan. Tämä on päätelty heuristisesti palvelutapahtumatunnuksen ja asiakirjojen aikaleimojen perusteella.



### Sairastavuus

Sairastavuusanalyysien helpottamiseksi on luotu myös taulu A\_COMORBIDITY, joka sisältää tiedot kaikista niistä sairauksista, joita PTA ICD-10-diagnoosin perusteella poimittu henkilö sairastaa. Taulun sarakkeet ovat seuraavat:

Nimi	Tietotyyppi	Selite
pid	raw(16)	Potilaan pseudonymisoitu tunniste
disid	integer	Sairauden tai toimenpiteen numero, viite dimensioon d_disease; ks. yllä oleva taulukko
disdate	date	Varhaisin päivämäärä, jolloin sairaudesta on olemassa diagnoosi tai toimenpidekirjaus
dgid	guid	Viite tauluun A_DIAGNOOSI, jos sairaus perustuu ICD-10-luokitukseen
procid	guid	Viite tauluun A_TOIMENPIDE, jos sairaus perustuu tehtyyn toimenpiteeseen

### Lääkitys

Taulu A\_MED on yksinkertainen taulu, joka sisältää yhteenvedon potilaan lääkityksestä perustuen Reseptikeskuksesta saatuun aineistoon. Sen sarakkeita ovat mm. seuraavat:

Nimi	Tietotyyppi	Selite
pid	raw(16)	Potilaan pseudonymisoitu tunniste
medypeid	integer	Lääkeryhmän tunnistenumero, ks. taulukko alla
prid	raw(16)	Viite lääkemääräykseen so. tauluun A_PRESCRIPTION so. ko. taulun CDA_ID-avaimeen
dlvcount	integer	Ko. lääkemääräyksen perusteella tehtyjen lääkeostojen lukumäärä
firstdlv	date	Ensimmäisen oston päivämäärä
latestdlv	date	Viimeisimmän oston päivämäärä

Lääkeryhmien tunnisteen (medypeid, dimensio *d\_medtype*) ovat tässä taulussa seuraavat:

medypeid	ATC	Selite
1	A10	Diabeteslääkitys
2	C02-C09	Verenpainelääkitys
3	C09	Reniini-angiotensiinijärjestelmään vaikuttavat lääkkeet
4	B01A	Veren hyytymiseen vaikuttava lääkitys
5	C10	Lipidilääkitys

Huomaa, että taulussa A\_PATIENT oleva sarake *medtypeid* on hienojakoisempi diabeteslääkityksen jaottelu:

medtypeid	parentid	ATC	Selite
10	10	A10B ei A10A	<b>Tabletti- tai GLP1-hoitoinen diabetes</b>
20	20		<b>Yhdistelmähoitoinen diabetes</b>
21	20	A10B ja ( A10AC tai A10AE)	– perusinsuliini
22	20	A10B ja A10AB ja (A10AC tai A10AE)	– monipistoshoido
23	20	A10B ja A10AB	– ateriainsuliini
24	20	A10B ja A10AD	– yhdistelmänsuliini
30	30		<b>Insuliinidiabetes</b>
31	30	A10A ja (A10AC tai A10AE)	– vain perusinsuliini
32	30	A10A ja A10AD	– vain ateriainsuliini
33	30	A10A ja A10AB ja (A10AC tai A10AE)	– perus- ja ateriainsuliini
34	30	A10A ja A10AD	– vain sekoiteinsuliini
40	40		<b>Lääkkeetön hoito tai elämäntapahoito</b>

### Laboratoriotulokset

Laboratorionimikkeet oli tässä pilotissa valittu vain Kuntaliiton laboratoriotutkimusnimikkeistä ("Laboratoriotutkimusnimikeld") seuraavasti:

Tutkimus-lyhenne	Laboratoriotutkimusnimike	Tutkimustulos-yksikkö	Laboratoriotutkimusnimikeld
B -HbA1c	B -Hemoglobiini-A1c	mmol/mol	6128
cU-Alb	cU-Albumiini	ug/min	3557
fB-Leuk	fB-Leukosyytit	umol/l	2218
fP-Kol-LDL	fP-kolesteroli, low density lipoproteiinit	mmol/l	4599
P -Krea	P -Kreatiniini	umol/l	4600
Pt-GFRe-CG	Pt-Glomerulussuodosnopeus, estimoitu, Cockcroftin-Gaultin kaava	ml/min/A	4999
Pt-GFRe-MD	Pt-Glomerulussuodosnopeus, estimoitu, MDRD-tutkimuksen kaava (mittayksikkö: ml/min/1.73 m <sup>2</sup> )	ml/min/1.7	6000
U -AlbKre	U -Albumiinin ja kreatiniinin suhde	mg/mmol	4511

Näiden lisäksi laskettiin Pt-GFReEPI (mittayksikkö: ml/min/1.73 m<sup>2</sup>) CKP-EPI-laskentayhtälöiden avulla:

Naiset	P -Krea ≤ 62	$GFRe = 144 \times (P -Krea/61,9)^{-0,329} \times 0,993^{ikä}$
	P -Krea > 62	$GFRe = 144 \times (P -Krea/61,9)^{-1,209} \times 0,993^{ikä}$
Miehet	P -Krea ≤ 80	$GFRe = 141 \times (P -Krea/79,6)^{-0,411} \times 0,993^{ikä}$
	P -Krea > 80	$GFRe = 141 \times (P -Krea/79,6)^{-1,209} \times 0,993^{ikä}$

Taulu A\_LAB kokoaa laboratoriotulokset yhteen helpottaen laboratorioanalyysien tekoa. Taulun sarakkeita ovat:

Nimi	Tietotyyppi	Selite
labid	guid	Analyysin yksikäsitteinen generoitu surrogaattivain
pid	raw(16)	Potilaan pseudonymisoitu tunniste
asiakirjaoid	oid	Laboratoriovastaus-asiakirjan tunniste
merkintaoid	oid	Laboratoriomerkinän tunniste
entryoid	oid	Laboratoriotutkimuksen tunniste; kolmikko asiakirjaoid, merkintaoid, entryoid toimii viitteenä faktatauluun F_LABORATORIOVASTAUS
laboratoriotutkimusnimike	varchar2(100)	Laboratoriotutkimuksen virallinen, Kuntaliiton hyväksymä nimi
tutkimuslyhenne	varchar2(100)	Laboratoriotutkimuksen virallinen, Kuntaliiton hyväksymä nimilyhenne; viite tauluun D_LAB_UNITS
tutkimustulosarvo	varchar2(30)	Tutkimuksen tuloksen arvo
tutkinustulosyksikko	varchar2(30)	Tuloksen yksikkö
tutkimustulosteksti	nvarchar2(1024)	Tutkimuksen tuloksen mahdollinen tekstimuotoinen esitys
tutkimusaika	timestamp	Tutkimustuloksen valmistumisajankohta
value	number	Tuloksen numeerinen arvo

**Fysiologiset mittaukset**

Pilottiaineistoon poimittiin seuraavat fysiologiset mittaukset:

<b>Id</b>	<b>Fysiologinen mittaus</b>
8480-6	Systolinen verenpaine
8462-4	Diastolinen verenpaine
29463-7	Paino
8302-2	Pituus
59574-4	BMI, painoindeksi
8280-0	Vyötärön ympärys
8478-0	Keskiverenpaine
8287-5	Pään ympärys
8591-0	Keskuslaskimopaine
41982-0	Kehon rasvaprosentti
59575-1	Painoindeksi, lapset
3138-5	Pituus, ilmoitettu
3141-9	Paino, mitattu
3142-7	Paino, ilmoitettu
8138-5	Pituus, ilmoitettu tai arvioitu
8137-7	Pituus, mitattu
8453-3	Diastolinen verenpaine istuen
8459-0	Systolinen verenpaine istuen

Analyytikannan taulu *A\_FysiologinenMittaus* sisältää PTA-diagnosoitujen diabeetikoiden kaikki arkistoon rakenteisesti kirjatut ko. suureiden mittausarvot. Seuraavassa keskeisimmät taulun sarakkeet:

<b>Nimi</b>	<b>Tietotyyppi</b>	<b>Selite</b>
fid	raw	Mittauksen yksikäsitteinen GUID-tunniste
pid	raw	Potilaan pseudonymisoitu tunniste
fysiologinenmittausid	nvarchar2(50 char)	Mittauksen tunnistekoodi
fysiologinenmittaus	nvarchar2(255 char)	Mittauksen tunnisteen suomenkielinen selite
mittaasteikko	nvarchar2(100 char)	Mittauksen mitta-asteikko, pitkä versio
mittausaika	timestamp(6)	Mittausajankohta (päivän tarkkuudella)
mittaustulosnumeerinen	number(10,2)	Tulosarvo numeerisena

Nimi	Tietotyyppi	Selite
mittaustulosyksikko	nvarchar2(50 char)	Mittausyksikön lyhenne
mittaustulos-torstina	nvarchar2(255 char)	Tulos tekstinä, jos ei numeerinen
asiakirjaoid	nvarchar2(100 char)	Potilasasiakirjan OID-tunniste
merkintaoid	nvarchar2(100 char)	Merkinnän OID-tunniste
entryoid	nvarchar2(100 char)	Merkintärivin OID-tunniste: asiakirjaoid, merkintaoid ja entryoid muodostavat mittauksen luonnollisen avaimen

Fysiologisten mittausten vähäisistä tallennusmääristä johtuen ei edetty niiden analyysiin.

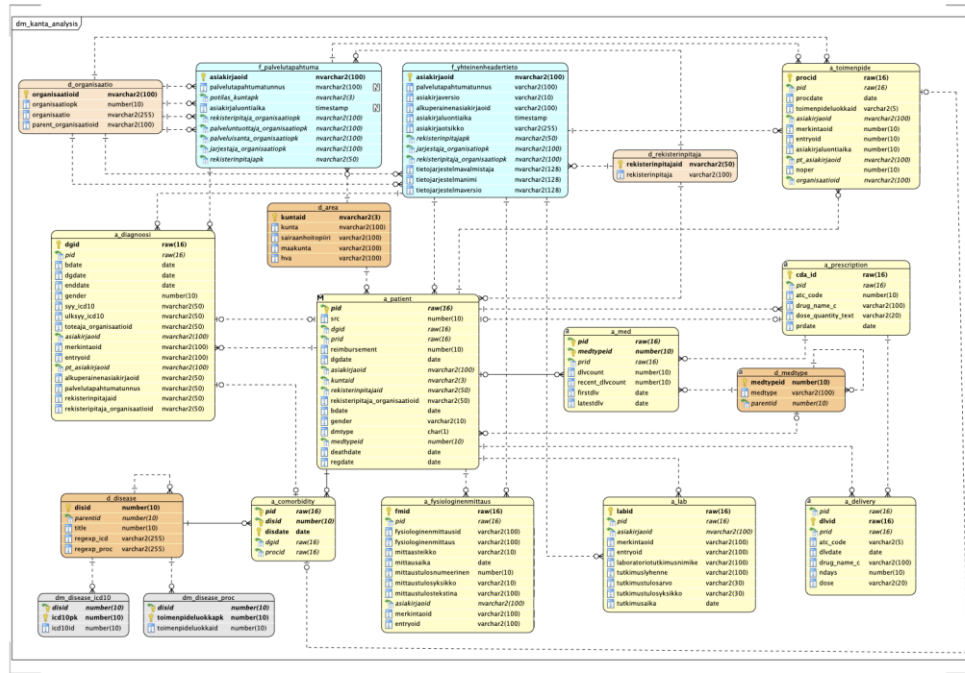
## Väestötiedot

Käytettävissä oli DVV:n väestörekisterin kopio, joka sisälsi rekisterin tilanteen sellaisena, kuin se oli 12.3.2021. Kopio on alun alkaen luotu 8.2.2014 elossa olevista, suomalaisen henkilötunnuksen omaavista henkilöistä. Tätä kopiota on sittemmin päivitetty säännöllisesti syntymä-, muutto- ja kuolintapahtumilla. Kopion syntyhistoria tarkoittaa sitä, että siitä puuttuvat kaikki ennen 8.2.2014 kuolleet henkilöt. Se ei sisällä myöskään niitä henkilöitä, joilla on väliaikainen henkilötunnus.

Rekisterikopiosta päivitimme tietokantaan potilaiden sukupuolen, syntymäajan, viimeisimmän kotikunnan sekä mahdollisen kuolinajan.

## ER-kaavio

Seuraava ER-kaavio havainnollistaa edellä kuvattujen taulujen keskinäisiä suhteita:



Kaaviossa entiteettien taustavärillä on seuraavat merkitykset:

- sininen kuvaa alkuperäisiä Kanta-faktatauluja (vrt. [teknologia-liitteen Kanta-ER-malliin](#))
- ruskea kuvaa dimensiotauluja (vaaleammat ovat alkuperäisiä Kanta-dimensiotauluja)
- harmaa muita metatauluja
- keltainen analyysitauluja

### Nimeämiskäytänteistä

Taulujen ja sarakkeiden nimet ovat englanninkielisiä silloin, kun niillä ei ole suoraa vastinetta alkuperäisessä tähtimallisessä Kanta-skeemassa. Suomenkieliset nimet pyrkivät olemaan täsmälleen samoja kuin Kanta-vastineensa.

Taulujen etuliitteillä on pyritty ryhmittelemään tauluja käyttötarkoituksensa ja EARP (entry, analysis, reporting, publish) -linkaarimallin mukaan:

- f\_ viittaa alkuperäisiin Kanta-faktatauluihin
- b\_ viittaa alkuperäisiin Kanta-siltatauluihin (bridge)
- d\_ viittaa dimensiotauluihin, jotka voivat olla alkuperäisiä Kanta-dimensioita tai analyysikantaa varten luotuja
- dm\_ viittaa diabetesrekisterin muihin metatietotauluihin
- e\_ viittaa ns. entry-tason tauluihin (THL:n EARP-mallissa) – tässä skeemassa näitä ei juuri esiinny, koska alkuperäiset f\_ ja b\_-taulut oikeastaan edustavat niitä
- a\_ viittaa analyysitauluihin (kuten "A" eArp:ssa)
- r\_ viittaa raportointitason tauluihin, jotka sisältävät yleensä jo aggregoitua dataa (eaRp); ovat esimerkiksi tilastokutiuita.

Esimerkiksi taulu a\_diagnoosi on analyysitaulu, jolle on käsitteellinen vastine Kanta-tietokannassa (f\_diagnoosi). Tämän vuoksi taulun nimi on suomeksi ja on yhteneväinen Kanta-

nimeämisen kanssa. Toisaalta a\_patient on analyysitaulu, jolle ei ole suoraa vastinetta Kanta-skeemassa. Sen vuoksi taulun nimi on englanninkielinen.

Kanta-skeeman nimeämiskäytänteitä ovat mm.

- dimensioiden primaariavaimet päättyvät merkkijonoon "pk" (primary key); faktataulujen dimensioihin viittaavat vierasavaimet on nimetty vastaavalla tavalla;
- dimensioissa esiintyvät koodiarvot päättyvät "id"-loppuun
- dimensioiden selitysteksti on samanniminen kuin dimensiotaulun nimikin
- oid-tyyppisten sarakkeiden nimet päättyvät "oid", esimerkiksi "asiakirjaoid", "merkintaoid", "entryoid" ja "palvelutapahtumaoid".

Esimerkiksi dimensiotaulun d\_icd10 primaariavain on "icd10pk", icd-koodiarvon sarake "icd10id" ja diagnoosin selitysteksti "icd10".

Huomioitavaa: Tietyssä vaiheessa talvella 2021 Kela päätti luopua surrogaattityyppisten viiteavaimien käytöstä. Sen jälkeen faktataulujen "pk"-sarakkeet sisälsivätkin kyseisen dimension ns. luonnollisen avaimen, esimerkiksi diagnoosin\_syyicd10pk viittasikin nyt d\_icd10-taulun sarakkeeseen icd10id eikä icd10pk. Muutoksen tarkoitus oli ratkaista se ongelma, että monet, dimensioihin kirjaamattomat koodiarvot kääntyivät latausvaiheessa arvoksi "Tunte-maton".

## Validoinnin tukiaineistot

### Hilmo-tiedot

Laskentaympäristöön kopioitiin THL:n Hoitoilmoitusjärjestelmästä sekä TerveysHilmon (erikoissairaanhoidon ESH) että AvoHilmon (perusterveydenhuolto PTH) keskeinen aineisto viime vuosilta. Potilaiden henkilötunnisteet pseudonymisoitiin samalla algoritmilla kuin edellä kuvatussa perusaineistossa, mikä mahdollisti yksilökohtaisen vertailun.

### TerveysHilmo (ESH)

Taulu *thl.tireka\_hilmo\_iso* sisälsi seuraavat sarakkeet:

Nimi	Tietotyyppi	Selite
isoid	number	Yksikäsitteinen avaintunniste, primaariavain
hetu_aes	raw	Pseudonymisoitu henkilötunniste
sukup	varchar2(1 char)	Sukupuolikoodi (1=mies, 2=nainen)
syntv	number	Syntymävuosi
koku	varchar2(3 char)	Kotikunnan koodi
ikat	number	Ikä vuosina
tupva	date	Tulopäivä sairaalaan
lpvm	date	Lähtöpäivä
toipvm	date	Toimenpiteen päivämäärä
pala	varchar2(2 char)	Palvelualakoodi
ea	varchar2(3 char)	Erikoisalakoodi
jatkoh	varchar2(2 char)	Jatkohoitokoodi
ilaji	varchar2(1 char)	Ilmoituslajin koodi
paltu	varchar2(5 char)	Palveluntuottajan TOPI-koodi
paltutar	varchar2(5 char)	Palveluntuottajan tarkenninkoodi
oid	varchar2(55 char)	Palveluntarjoajaorganisaation OID
tiedosto_id	number(38,0)	Lähdetiedoston tunnistenumero
lataus_pvm	date	Lähdetiedoston latauspäivämäärä



Taulu *thl.tireka\_hilmo\_iso\_icd10* sisältää hoitojaksoilla kirjatut ICD-10-diagnoosit:

Nimi	Tietotyyppi	Selite
isoid	number(38,0)	Hoitojakson yksikäsitteinen tunniste, vierasavain
kentta	varchar2(30 char)	Kuvaa diagnoosin tyyppiä (pää-, sivu-, pikäaikaisdiagnoosi, ulkoinen syy, tapaturmatyyppi tms.)
n	number(38,0)	Diagnoosin järjestysnumero alkaen nolasta
koodi	varchar2(5 char)	ICD-10-koodi, ilman pistettä tai muita erityismerkkejä

Vastaavasti taulu *thl.tireka\_hilmo\_iso\_tmp* sisältää kullakin hoitojaksolla kirjatut toimenpiteet:

Nimi	Tietotyyppi	Selite
isoid	number	Hoitojakson yksikäsitteinen tunniste, vierasavain
tmpkentta	varchar2(9 char)	Toimenpidettä kuvaava tunniste: pää- vai muu toimenpide, ensisijaisuus
tmpkoodi	varchar2(5 char)	THL Toimenpideluokituksen toimenpidekoodi (pohjana aikanaan ns. Nomesco-luokitus)

Lisäksi oli käytettävissä taulu *thl.hilmo\_iso\_meta*, joka sisältää kunkin ilmoituksen metatiedot kuten ilmoituksen ensimmäisen ja viimeisimmän version lähetyspäivät ja mitkä ovat tiedot lähettäneen ohjelman toimittaja ja nimi.

### AvoHilmo (PTH)

Perusterveyshuollon hoitoilmoitukset on kirjattu tauluun *thl.tireka\_avo\_palvelutapahtuma*, jonka sarakkeet ovat seuraavat:

Nimi	Tietotyyppi	Selite
tapahtuma_tunnus	number	Yksikäsitteinen palvelutapahtuman tunnus, primariivain
hetu_aes	raw	Asiakkaan pseudonymisoitu henkilötunniste
asiakas_sukupuoli	number	Asiakkaan sukupuoli, 1=mies, 2=nainen
asiakas_syntyma aika	date	Asiakkaan syntymäaika
asiakas_kotikunta	varchar2(3 char)	Asiakkaan kotikunnan koodi
paino	number	Asiakkaan paino
pituus	number	Asiakkaan pituus
tupakointi	number	Asiakkaan tupakointi
kaynti_alkoi	timestamp(6)	Käynnin alkamisajankohta
kaynti_loppui	timestamp(6)	Käynnin loppuminen
peruutus_ajankohta	timestamp(6)	Käynnin peruutusajankohta
palveluntuottaja	varchar2(5 char)	Palveluntuottajan TOPI-koodi

palveluntuottaja_yksikko	varchar2(255 char)	Palveluntuottajan OID-koodi
paino_yksikko	varchar2(3 char)	Henkilön painon yksikkö (null tai G tai KG)
lahetys	number	Lähetysosan tunnistenumero
seurantatietue_paivitetty	timestamp(6)	Tietueen päivityksen ajankohta

Lähetysosan numeron perusteella voidaan selvittää omasta taulustaan (*thl.tireka\_avo\_lahetys*) mm. milloin sanoma on vastaanotettu ja mikä oli potilastietojärjestelmän nimi ja sen toimittaja.

Kuhunkin palvelutapahtumaan voi liittyä nolla tai useampia indikaatioita, jotka on kirjattu tauluihin *thl.tireka\_avo\_kayntisyys\_icd10* tai *thl.tireka\_avo\_kayntisyys\_icpc2* riippuen, mitä nomenklatuuria on käytetty koodauksessa käytetty.

## Fimlab-pilottiaineisto

Fimlabin toimittamat laboratoriotiedot on kirjattu tauluun *ks\_p\_labrat*, jonka sarakkeet ovat seuraavat:

Nimi	Tietotyyppi	Selite
pid	raw	Asiakkaan pseudonymisoitu henkilötunniste
naytteenotto_pvm	date	Laboratoriotutkimuksen näytteenottopäivä
tilaava_laitos	varchar2(128 char)	Tutkimuksen tilanneen laitoksen nimi
tutk_nro	number	Tutkimuksen tunnistekoodi (Kuntaliiton mukainen tai sisäinen)
tutk_lyhenne	varchar2(10 char)	Tutkimuksen lyhenne
tulos	varchar2(26 char)	Tutkimuksen tuloksen arvo
yksikko	varchar2(16 char)	Tuloksen mittayksikkö
viitearvoala	number	Mahdollisen viitearvon (tai mittaustarkkuuden?) alaraja
viitearvoyla	number	Mahdollisen viitearvon (tai mittaustarkkuuden?) yläaraja
viitearvoteksti	varchar2(16 char)	Mahdollinen viitearvo (tai mittaustarkkuus?) tekstinä
poikkeustilanneviesti	varchar2(20 char)	Mahdollinen poikkeustilanteesta tai virheestä kertova merkki
shp	varchar2(5 char)	Tutkimuksen tilanneen laitoksen sairaanhoitopiiri

## Siun Sote -pilottiaineisto

Siun Soten tällä hetkellä käytössä oleva aineisto sisältää tiedot laboratoriomittauksista:

- ss\_labrat ja ss\_pikalabrat: Sisältävät potilaiden laboratoriotutkimusten ja HbA1c-pikamittausten tulokset ja päivämäärät.

Lisäksi Siun Sotesta kerättiin vuoden 2019 alussa seuraavat tiedot Hilmo-tietojen kattavuuden ja paikkansapitävyyden validoimiseksi:

- ss\_henkilot: Sisältää listauksen kaikista Siun Soten diabetespotilaista ja heille henkilötiedot: ikä, sukupuoli, kuolinpäivä (jos kuollut), kotikunta ja diagnoosi, jonka mukaan heidät valittiin aineistoon.
- ss\_dg: Diagnoositaulu. Sisältää kaikki potilaiden saamat diagnoosit ja niiden viimeisimmän asettamispäivän
- ss\_esh ja ss\_pth: Erikoissairaanhoidon ja perusterveydenhuollon käynnit
- ss\_hoitojaksot: Erikoissairaanhoidon hoitojaksot.

HbA1c-pikamittausten ja tavallisten mittausten tulosten mahdollisesta erilaisuudesta oli huolta, mutta ainakin tässä aineistossa niiden tulokset vaikuttivat lyhyen tarkastelun perusteella samantapaisilta.

## Terveystalon työterveysasiakkaat

Terveystalolta kerättiin vuoden 2019 alussa seuraavat tiedot Hilmo-tietojen kattavuuden ja paikkansapitävyyden validoimiseksi: (ks lisätietoja [Terveystalo](#))

- tt\_henkilot: Lista TT:n työterveyshuollon diabetespotilaista ja heille sukupuoli ja syntymäaika.
- tt\_kaynnit: Kaikki diabetespotilaiden käynnit (päivämäärä, diagnoosi, erikoisala, toimipiste ja annetut lomapäivät).
- tt\_koodisto: Koodisto, nimet ja kuvaukset aineistossa käytetyille lyhenteille ja koodeille. (toimipaikka, labra jne.)
- tt\_laakitys: Potilaiden lääkkeet: ATC ja määräyspäivä.
- tt\_labra: Potilaiden laboratoriomittauksien tulokset ja -päivät.

## Julkisten rekisterinpitäjien kaikki PTA-laboratoriotulokset

Aineistoa analysoitaessa alkoi käydä ilmeiseksi, että monet rekisterinpitäjät ja näiden käyttämät ohjelmistot kirjaavat laboratoriotuloksia käyttäen ns. paikallisia tutkimusnimikkeitä [Kuntaliiton virallisen laboratoriotutkimusnimikkeistön](#) sijasta. Jotta saataisiin selvyttä ensinnäkin laboratoriomäärittysten volyymin rekisterinpitäjittäin ja toisaalta käytetyn nimikkeistön kirjavuudesta, luotiin seuraavanlainen aggregaattitaulu ("aggregoidut\_labrat") **kai-kista PTA:han rakenteisesti tallennetuista laboratorioanalyseistä:**

Nimi	Tietotyyppi	Selite
rekisterinpitäjät	varchar2(255 char)	Rekisterinpitäjän OID-tunniste, luokittelija
rekisterinpitäjä	varchar2(255 char)	Rekisterinpitäjän nimi, luokittelija
tietojärjestelmänimi	varchar2(255 char)	Tietojärjestelmän nimi, luokittelija
nimikeid	varchar2(255 char)	Laboratoriotutkimusnimikkeen koodi, luokittelija. Jos numero < 10000, niin KL:n koodi, muuten paikallisen tutkimusnimikkeen koodi
nimike	varchar2(255 char)	Paikallisesta tutkimusnimikkeestä käytetty lyhenne, luokittelija
dccount	number(19,0)	Em. luokittelijoiden määrittämien PTA-asiakirjaversioiden kokonaislukumäärä
setcount	number(19,0)	Em. luokittelijoiden määrittämien PTA-asiakirjojen (setID) lukumäärä
mindate	varchar2(255 char)	Ensimmäisen asiakirjan luontiaika
maxdate	varchar2(255 char)	Viimeisen asiakirjan luontiaika

Tästä aggregaattitaulusta luotiin vielä toinen versio, jossa rajauduttiin vuoden 2019 näytteisiin (aggregoidut\_labrat\_2019).

Taulun perusteella voidaan muodostaa käsitys mm. niiden analyysien määrästä, joissa nimikkeenä ei ole käytetty Kuntaliiton koodistoa, vaan jotain paikallista numerointia. Varsinaisesta pilottiaineistosta puuttuvat kaikkien niiden laboratorioanalyysien tulokset, joissa tunnisteenä on jokin muu kuin virallinen Kuntaliiton laboratorionimikkeistön koodi (esim. "6128" HbA1c-analyseille).

Tästä aggregaattitaulusta pyrittiin löytämään eri analyysien lukumäärät seuraavankaltaisilla säännöllisillä lausekkeilla:

```
-- P -Krea:
"nimikeid" = '4600' or regexp_like("nimike", 'p\s?-krea', 'i' )

-- B -HbA1c:
"nimikeid" = '6128' or ( regexp_like("nimike", 'hba1c\*?$', 'i' )
and not regexp_like( "nimike", 'pika|oma', 'i' ))
```

## Tilastolliset menetelmät

Ikävakioidut prevalenssit laskettiin suoralla ikävakioidinnalla, referenssipopulaationa koko Manner-Suomen väestön ikärakenne. Ikä oli tässä luokiteltu 1-vuotisikäryhmiin.

Lisäksi ikä- ja sukupuolivakioituja prevalenssisuhteita ja näiden 95 %:n luottamusvälejä estimoitiin yleistetyllä lineaarisella mallilla (GLM), jakaumana binomijakauma, käyttäen log-linkifunktiota ja painoina populaation kokoa.

Kuolleisuutta havainnollistettiin Kaplan–Meier-estimaattien avulla piirtämällä eloonjäämisfunktion  $S(t)$  komplementti  $100*(1-S(t))$ . Vakioidut riskitiheyssuhteet (hazard ratios) estimoitiin Coxin elinaikamallilla. Elinaikaseurannan alkuhetkeksi valittiin varhaisin sellainen päivämäärä, jolta potilaalla oli diabetesmerkintä missä tahansa käytössä olevista tietolähteissä. Seuranta katsottiin päättyneeksi 12.3.2021, jos henkilö ei ollut kuollut ennen sitä.

## 3 Tulokset

### Potilaiden tunnistaminen

#### Potilaspopulaation muodostaminen ja kattavuuden arviointi

Potilaspopulaation muodostamista arvioitiin tunnistamalla diabetespotilaita eri tietolähteistä ja vertaamalla eri tietolähteistä löytyviä potilaspopulaatioita toisiinsa. Käytetyt tietolähteet ja poimintakriteerit on lueteltu alla.

- Potilastiedon arkistosta (PTA) potilaat, joilla esiintyi pää- tai sivudiagnoosina diabetesdiagnoosi (ICD-10: E10\*, E11\* tai E12\*-E14\*). Käytännössä tietoja alkaa tulla PTA:oon kattavasti valtaosasta sairauskertomusjärjestelmiä vuodesta 2017 alkaen.
- Kelan erityiskorvausoikeusrekisteristä (Etuus) potilaat, jotka olivat saaneet vuoden 2010 jälkeen erityiskorvattavuusoikeuden: 103 tai 215. Vuonna 2017 tapahtuneen perus- ja erityiskorvattavuusoikeuden erottamisen myötä kaikille aiemmin 103-oikeuden saaneille kirjattiin myös 215-oikeus.
- Lääkitystietokannasta (E-Resepti) potilaat, joille oli joskus määrätty diabeteslääkkeitä (ATC-koodi A10\* ). Käytännössä tietoja lääkemääräyksistä alkaa löytyä kattavasti vuodesta 2014 alkaen, ja eResepti tuli pakolliseksi 1.1.2017.
- TerveysHilmo ja Avohilmo-rekistereistä (Hilmo) potilaat, joilla esiintyi pää- tai sivudiagnoosina diabetesdiagnoosi (ICD-10: E10\*, E11\* tai E12\*-E14\*). Pilotissa oli käytettävissä TerveysHilmon tiedot 2017 alkaen ja Avoilmon tiedot 2020 alkaen.

Eri tietolähteiden potilaspopulaatioiden kattavuuden arvioimiseksi alkuperäisestä potilaspopulaatiosta rajattiin pois ICPC2-diagnoosin saaneet sekä raskausdiabeetikot, koska ICPC2- ja raskausdiabetesdiagnooseilla löytyy potilaita, joilla on vasta diabetesepäily tai alkava diabetes, joihin ei tehdä lääkemääräyksiä eikä myönnetä diabeteslääkitykseen korvattavuutta.

Kun verrattiin Hilmosta ja PTA:sta diagnoosien poiminnan pohjalta löytyviä potilaita, havaittiin että PTA:sta löytyy potilaita diabetesdiagnoosikirjauksiin perustuen kattavammin kuin Hilmosta. Ainoastaan Hilmosta löytyi vain 0,7 % kaikkien neljän tietolähteen summapopulaatiosta (Kuvat 3.2 ja 3.3). Hilmon tietopoiminta ei siis tuonut merkittävää lisää populaatiokattavuuteen PTA:n lisäksi.

Kun verrattiin PTA:sta, lääkitystietokannasta ja Kelan erityiskorvausoikeusrekistereistä löytyviä potilaita havaittiin, että 60 % potilaista löytyy kaikista kolmesta tietolähteestä (kuva 3.1). Lääkemääräyksiä oli tehty tai etuuksia myönnetty 20 %:lle potilaista ilman, että heiltä löytyi PTA:sta diabetesdiagnoosikirjausta. Tämä viittaa siihen, että diagnoosien kirjaaminen tai siirtyminen potilastiedon arkistoon ei ole vielä riittävällä tasolla. Tavoitteena olisi, että 80 %:lle summapopulaation potilaista löytyisi diabetestyyppin määrittelevä diagnoosi PTA:sta.

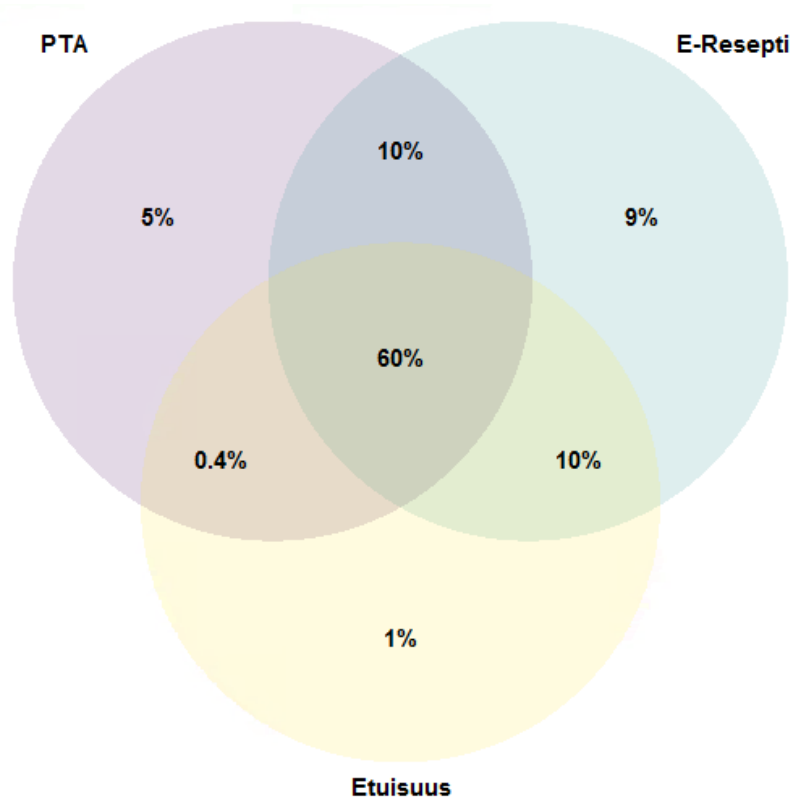
9 %:lle potilaista oli määrätty diabeteslääkettä, mutta ei kirjattu diabetesdiagnoosia tai myönnetty lääkekorvattavuutta. Selityksenä voi olla tilapäinen lääkkeen käyttö esimerkiksi kortisonihoidon tai raskausdiabeteksen vuoksi tai diabeteslääkkeiden käyttö muissa indikaatioissa, esimerkiksi lihavuuden vuoksi.

Vain 1 %:lla potilaista oli lääkekorvattavuus diabeteslääkkeisiin, mutta ei määrättyjä diabeteslääkkeitä tai kirjattua diagnoosia.

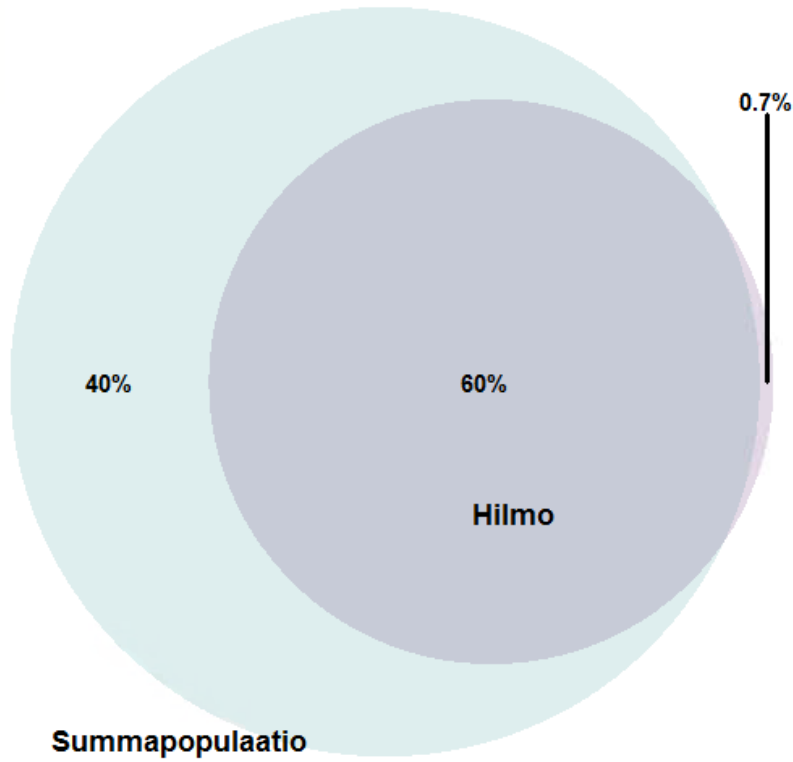
Koska mikään tietolähde ei yksiselitteisesti tunnista kaikkia diabeetikoita, päätettiin käyttää summapopulaatiota, johon poimitaan potilaat PTA:n diagnoosien, annettujen

lääkemääräysten ja myönnettyjen erityiskorvattavuusosoikeuksien perustella. Jatkossa on syytä raportoida diagnoosikattavuus ja sen vajeet hyvinvointialueittain. Lisäksi voidaan pyrkiä diabeteksen keston tunnistamiseen etsimällä jatkossa ensimmäinen diabeteksen ilmaantumisen merkki mistä tahansa kolmesta tietolähteestä.

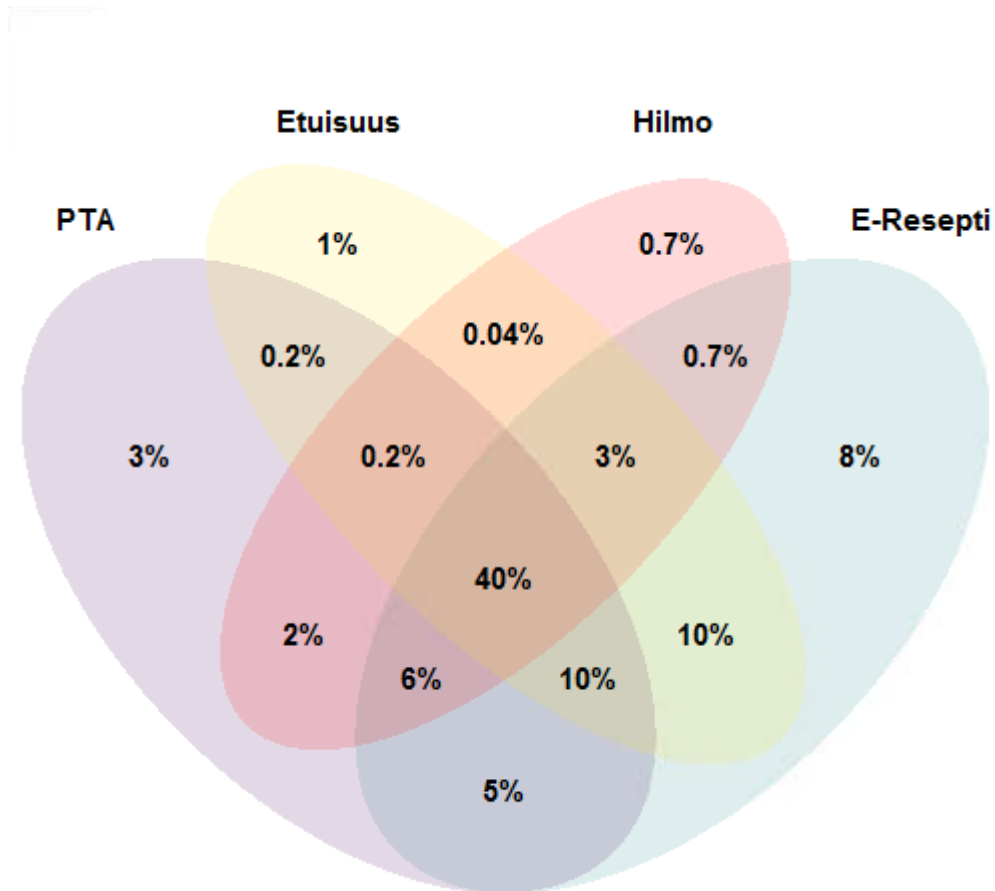
**Summapopulaatio = PTA:n diabetesdiagnoosit (ICD-10) + diabeteslääkemääräykset + myönnetyt erityiskorvattavuudet**



Kuva 3.1



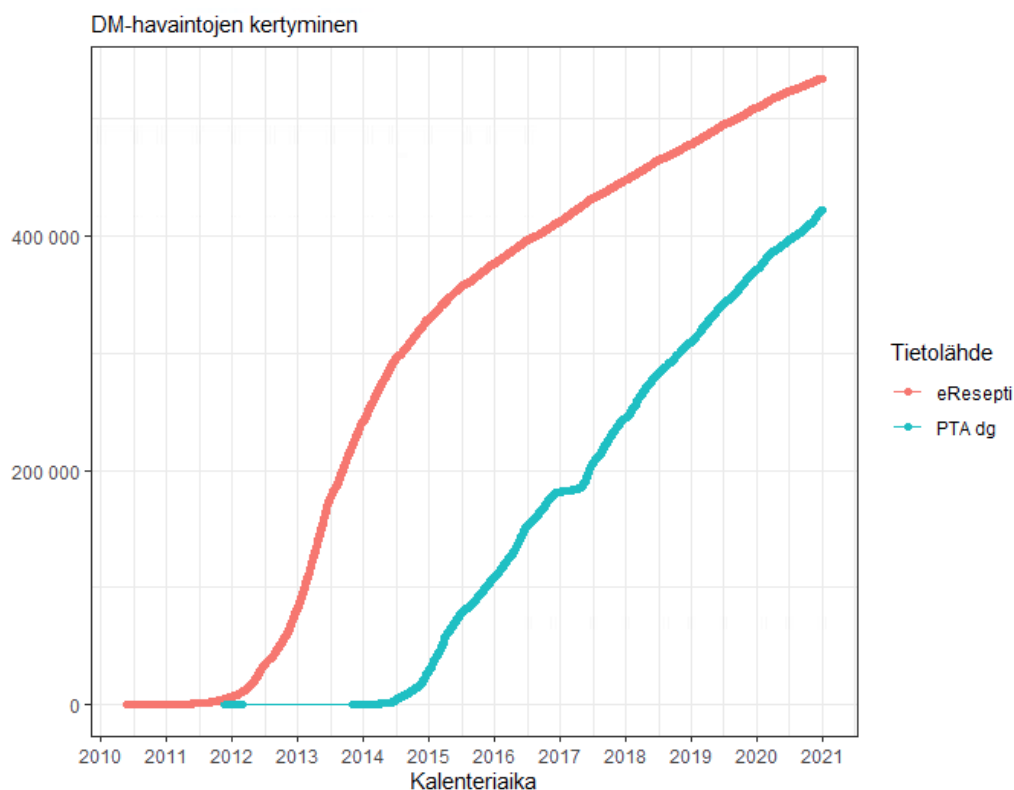
Kuva 3.2



Kuva 3.3



Seuraava DM-havaintojen kertymistä kuvaava kaavio havainnollistaa sen, kuinka sähköinen reseptijärjestelmä otettiin Suomessa käyttöön ennen potilastiedon arkistoa:



## Alueellinen prevalenssi

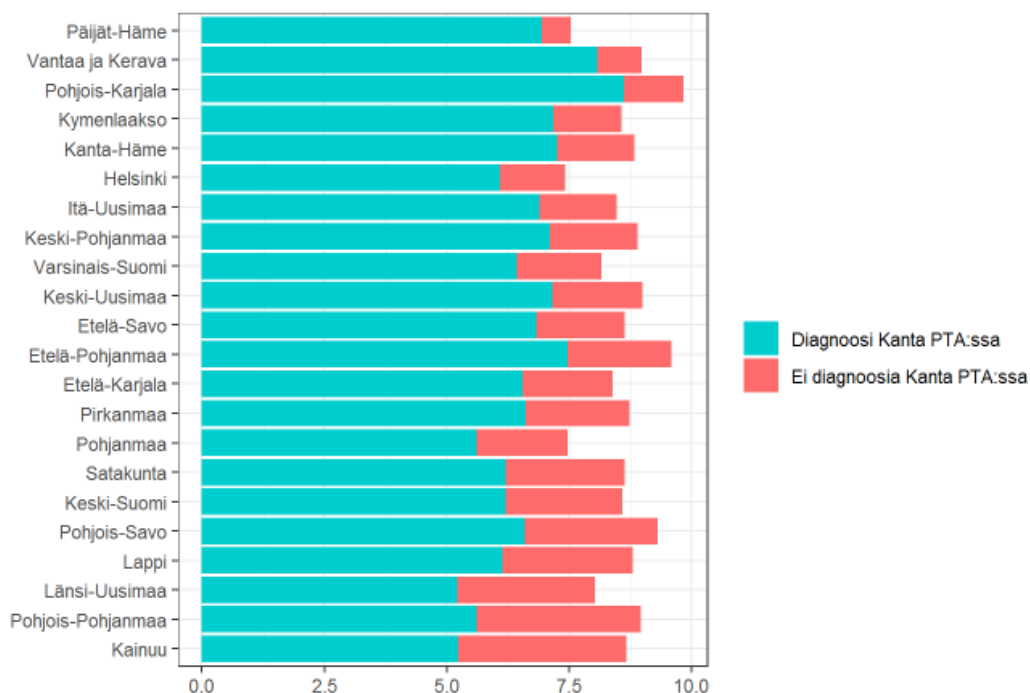
Kuvassa 3.4a näytetään PTA:n, lääkemääräysten ja etuuksien perusteella muodostettuun summapopulaatioon perustuva diabeteksen ikävakioitu prevalenssi hyvinvointialueittain. Diabeteksen prevalenssissa on eroja hyvinvointialueiden välillä.

Kuvassa 3.4a turkoosi palkki kuvaa PTA:sta diagnoosikirjausten perusteella löytyvän populaation osuutta summapopulaatiosta eli diagnoosikattavuutta hyvinvointialueittain. Mukana analysissa ovat vain elossa olevat diabeetikot, jotka löytyvät väestötietojärjestelmästä (VTJ). Aluetiedot perustuvat henkilön kotikuntaan (VTJ). Osalle henkilöistä ei löytynyt kotikuntaa Suomesta, joten heidät on jätetty pois alueittaisista tarkasteluista.

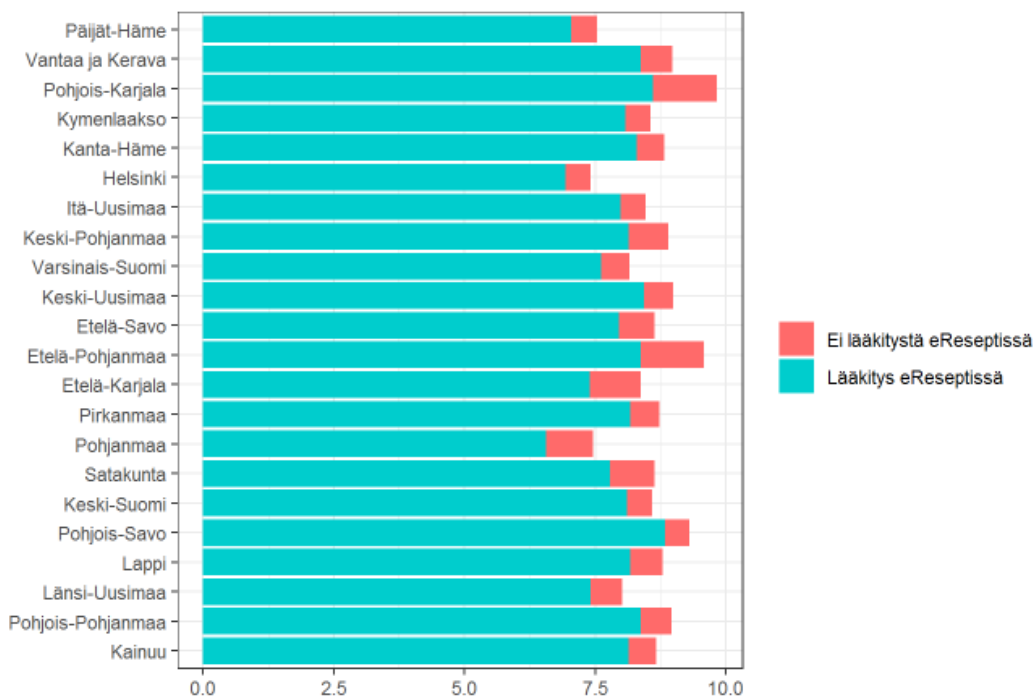
Hyvinvointialueiden välillä on vaihtelua diagnoosien kirjaamiskattavuudessa. Koska diabetestyyppi (tyyppi 1, tyyppi 2 ja muu) on määritetty PTA:n aineistosta löytyvän diagnoosin perusteella, tavoitellaan jatkossa sitä, että diabetesdiagnoosien kattavuus on yli 80 % kaikilla hyvinvointialueilla. Tässä vaiheessa diagnoosikattavuus on yli 80 % vain 8 hyvinvointialueella: Päijät-Hämeessä, Vantaalla ja Keravalla, Pohjois-Karjalassa, Kymenlaaksossa, Kanta-Hämeessä, Helsingissä, Itä-Uusimaalla ja Keski-Pohjanmaalla. Koska sairauskertomusjärjestelmät on kytketty eri aikoina potilastiedon arkistoon, voi osa kirjaamiskattavuuden mataluudesta selittyä teknisellä syyllä etenkin niiden järjestelmien kohdalla, jotka ovat liittyneet Potilastiedon arkistoon viimeisten joukossa. Selkeää systeemivirhettä matalaan PTA-diagnoosikattavuuteen ei tässä vaiheessa tunnistettu.

Kuvassa 3.4b näytetään vastaavasti ikävakioitu prevalenssi 100 asukasta kohden hyvinvointialueittain, ja turkoosi palkki kuvaa niiden potilaiden osuutta summapopulaatiosta, joille on määrätty diabeteslääkettä. Kuvasta voidaan havaita, että hyvinvointialueilla, joissa on matala diagnoosikattavuus, on kuitenkin korkea lääkekattavuus. Lääkekattavuudessa ei ollut

merkittävää vaihtelua hyvinvointialueiden välillä. Tämä viittaa siihen, että diabeetikot on tunnistettu ja heitä hoidetaan aktiivisesti, mutta diagnosoikirjaukset eivät jostain syystä kaikilta hyvinvointialueilta siirry PTA:n.



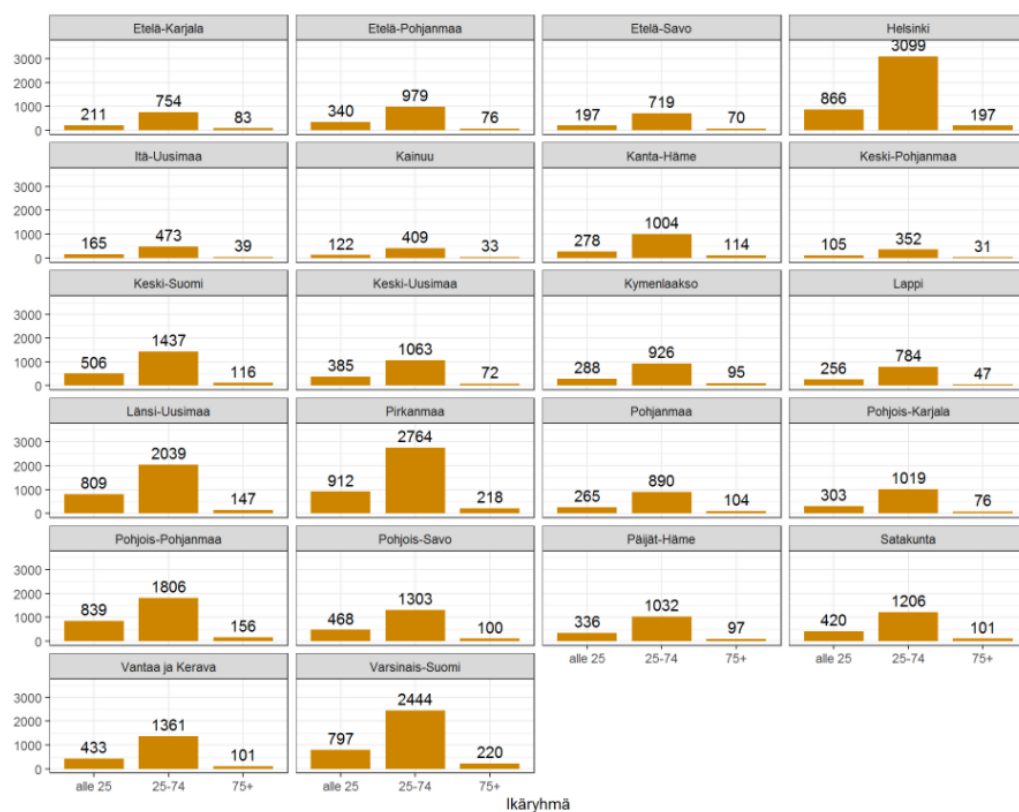
Kuva 3.4a: Ikävakioitu prevalenssi 100 asukasta kohden hyvinvointialueittain. Hyvinvointialueiden järjestys parhaimmasta PTA-diagnosikattavuudesta huonoimpaan.



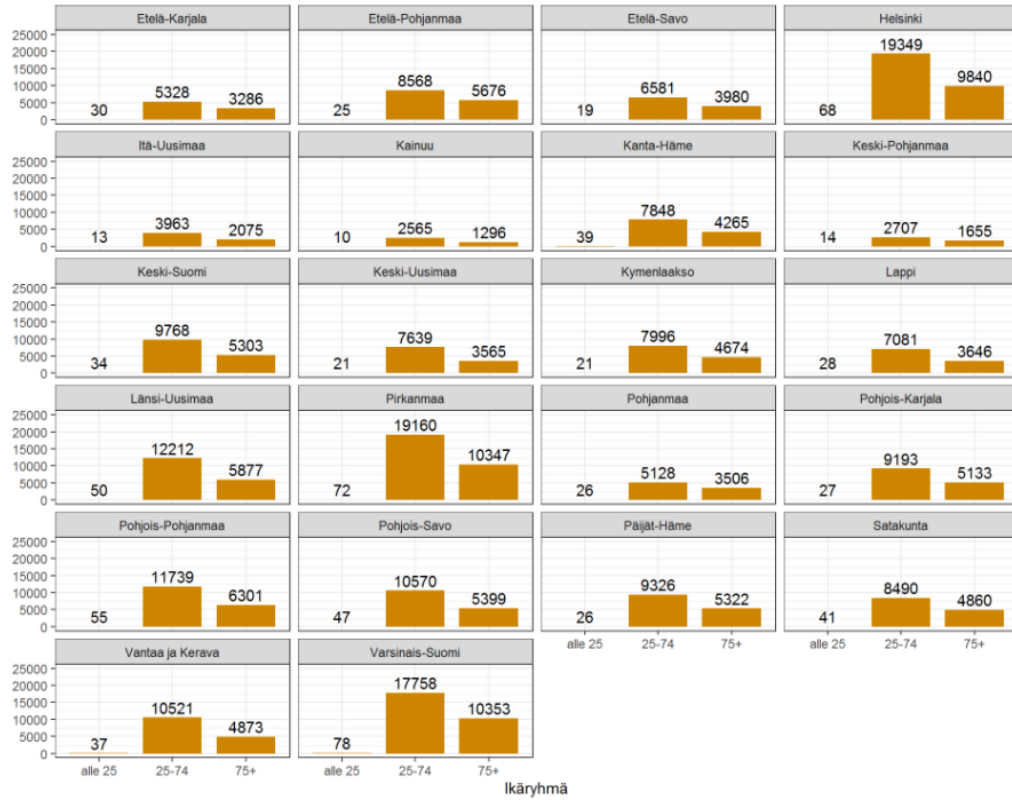
Kuva 3.4b: Ikävakioitu prevalenssi 100 asukasta kohden hyvinvointialueittain sen perusteella, löytyykö henkilölle diabeteslääkeresepti eReseptistä. Hyvinvointialueiden järjestys parhaimmasta PTA-diagnosikattavuudesta huonoimpaan.

## Esimerkkejä diabeteslaaturekisteriin tuotettavista diabeteksen prevalenssia ja diabeetikoiden lukumäärää kuvaavista raporteista

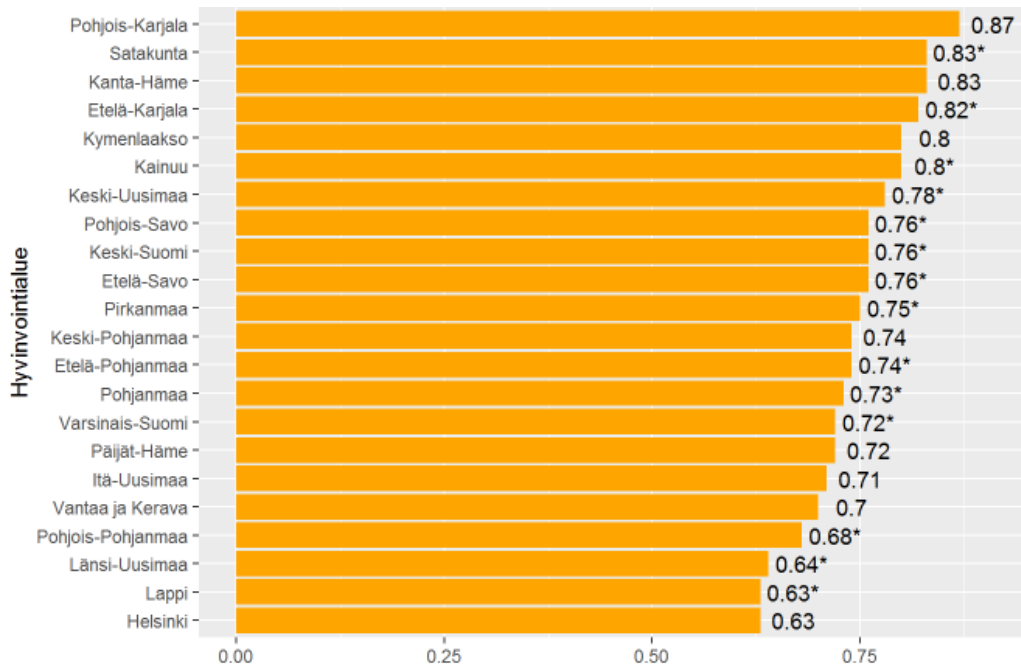
Niillä hyvinvointialueilla, joilla diagnoosien kirjauskattavuus on yli 80 %, voidaan diabeetikoiden lukumäärä ja ikävakioitu prevalenssi raportoida jatkossa ikä- ja diabetestyyppittäin (kuvat 3.5–3.8). Hyvinvointialueiden välillä on huomattavaa vaihtelua diabeetikoiden absoluuttisessa määrässä ikäryhmittäin (kuvat 3.4 ja 3.5) ja diabetestyyppien esiintyvyydessä (kuvat 3.7 ja 3.8). Esimerkiksi Helsingissä on 3099 25–74-vuotiaasta tyyppin 1 diabeetikkoa, kun Pohjois-Karjalassa heitä on 1019 (Kuva 3.5). Pohjois-Karjalassa tyyppin 1 ikävakioitu esiintyvyys on kuitenkin huomattavasti suurempi kuin Helsingissä (0,87 vs. 0,63). Toistaiseksi esimerkkitaulukoiden tuloksiin tulee suhtautua varauksella etenkin niiden hyvinvointialueiden kohdalla, joiden diagnoosien kirjaamiskattavuus on alle 80 % (muualla kuin Päijät-Hämeessä, Vantaalla ja Keravalla, Pohjois-Karjalassa, Kymenlaaksossa, Kanta-Hämeessä, Helsingissä, Itä-Uusimaalla ja Keski-Pohjanmaalla). Jos diagnoosikattavuus on alle 80 %, syntyy alla oleviin raporteihin valikoitumisharha, koska ilman diabetesdiagnoosia jääviä potilaita ei käytetä raporteissa.



Kuva 3.5: Hyvinvointialueiden tyyppin 1 diabeetikoiden lukumäärät ikäryhmittäin.

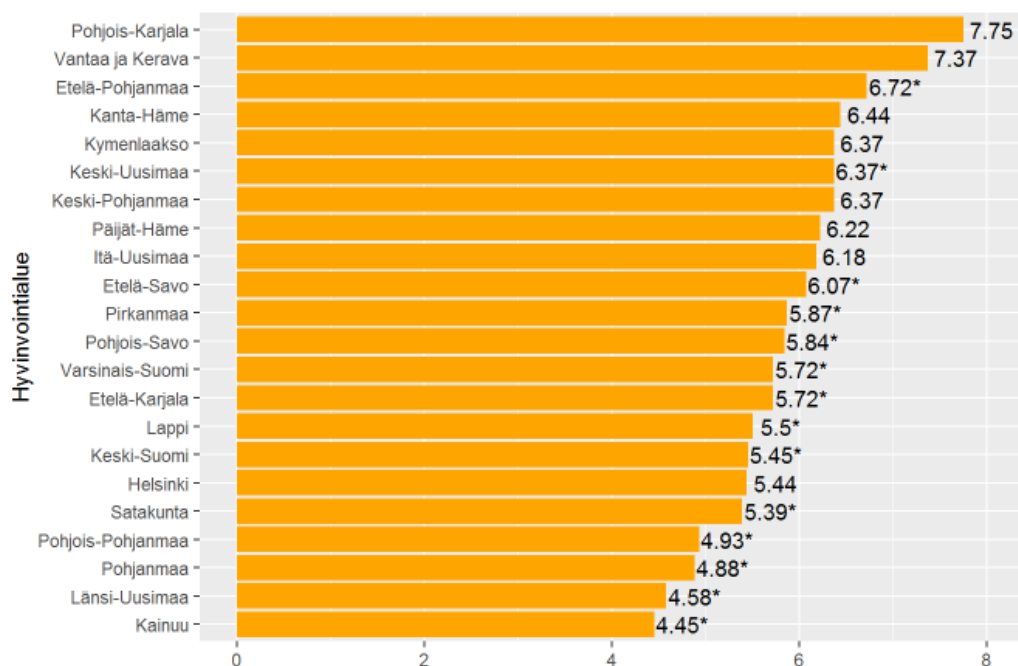


Kuva 3.6: Hyvinvointialueiden tyyppin 2 tai muu diabeetikoiden lukumäärät ikäryhmittäin.



Kuva 3.7: Tyyppin 1 ikävakioitu prevalenssi 100 asukasta kohden hyvinvointialueittain.

\*PTA:sta löytyneiden diagnoosien kattavuus alle 80 %.



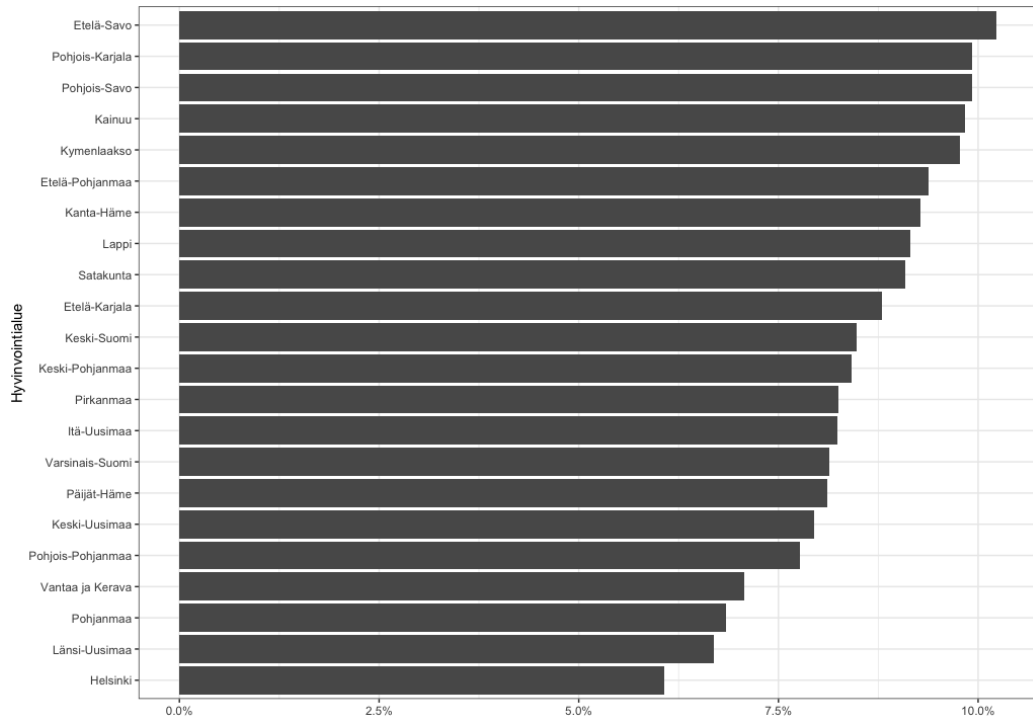
Kuva 3.8: Tyypin 2 ja muiden diabetestyyppien ikävakioitu prevalenssi 100 asukasta kohden hyvinvointialueittain.

\*PTA:sta löytyneiden diagnoosien kattavuus alle 80 %

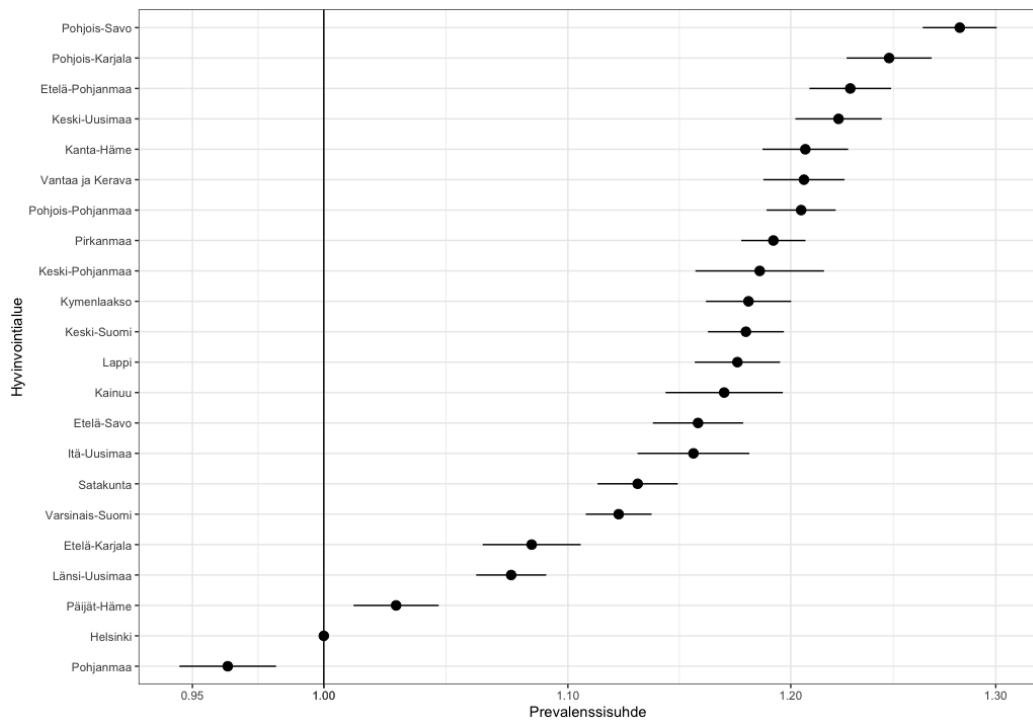
### Diabeteslääkemääräyksiin ja diagnooseihin perustuvan diabeteksen prevalenssin vertailu hyvinvointialueittain

Kuvassa 3.9 esitetään 31.12.2020 elossa olevien diabeteslääkemääräyksen saaneiden väestöosuuksia (prevalensseja) hyvinvointialueittain. Jos verrataan tätä kuvaa 3.4b-kuvaan, havaitaan, että pelkkää lääkemääräystä käytettäessä saadaan varsin samankaltainen tilannekuva hyvinvointialueiden välisistä eroista diabeteksen prevalenssissa kuin summapopulaatiota käytettäessä, joskin käyttämällä summapopulaatiota, diagnoosin perusteella tunnistetut diabeetikot nostavat esimerkiksi Etelä-Pohjanmaan korkeamman prevalenssin alueeksi kuin pelkkää lääketietoa käytettäessä.

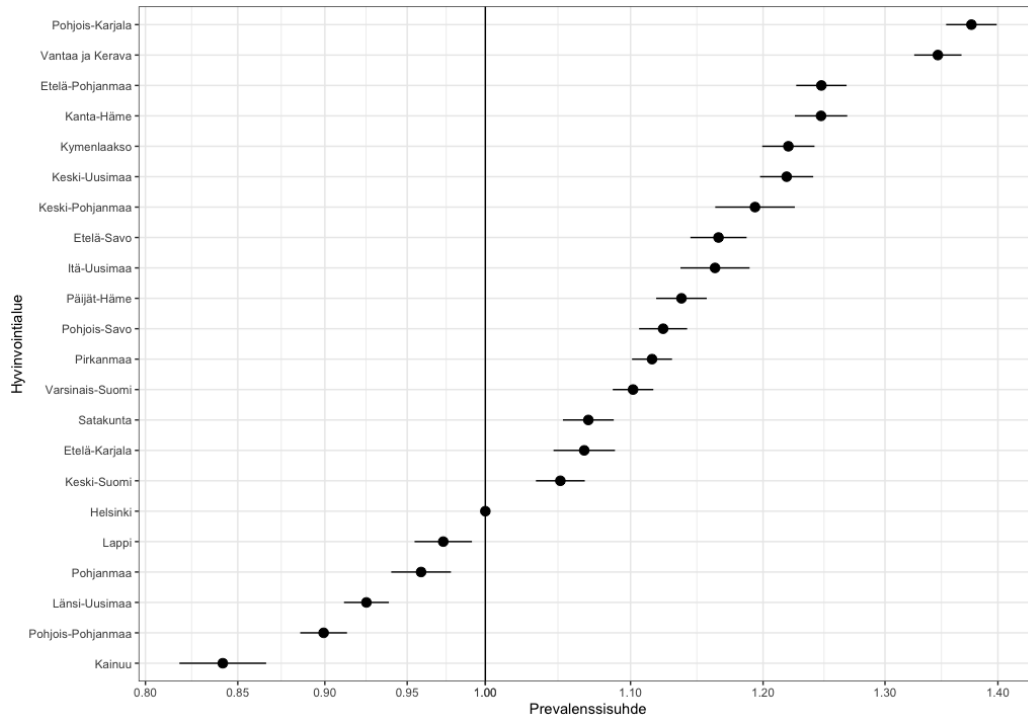
Jos lääkemääräyksiin perustuvia ikä- ja sukupuolivakioituja prevalensseja verrataan Helsingin prevalenssiin, havaitaan että Pohjanmaalla on matalampi ja kaikilla muilla hyvinvointialueilla korkeampi diabeteksen prevalenssi kuin Helsingissä (Kuva 3.10). Pelkästään diagnooseja käyttämällä vaikuttaa matalan diagnoosikattavuuden vuoksi virheellisesti siltä, että myös Kainuussa, Pohjois-Pohjanmaalla, Länsi-Uusimalla ja Lapissa olisi matalampi suhteellinen ikävakioitu prevalenssi kuin Helsingissä (Kuva 3.11). Tämä vahvistaa havaintoamme siitä, että diabeteksen prevalenssia ei tule arvioida pelkästään PTA:sta löytyviin diagnooseihin perustuen ja diagnoosien kattavuutta tulee pyrkiä parantamaan jatkuvalla kehittämistyöllä ja raportoimalla diagnoosikattavuuden vajeet.



Kuva 3.9. 31.12.2020 elossa olevien diabeteslääkemääräyksen saaneiden potilaiden väestöosuudet (prevalenssit) hyvinvointialueittain.



Kuva 3.10. Lääkemääräystietoon perustuvat ikä- ja sukupuolivakioidut prevalenssien suhteet hyvinvointialueittain verrattuna Helsingin prevalenssiin. Aineistossa ovat mukana ne ATC A10-lääkitystä saaneet potilaat, jotka olivat elossa 31.12.2020.



Kuva 3.11. PTA:sta löytyvään diabetesdiagnoositietoon perustuvat ikä- ja sukupuolivakioidut prevalenssien suhteet hyvinvointialueittain verrattuna Helsingin prevalenssiin. Aineistossa ovat mukana ne diabetesdiagnoosin saaneet potilaat, jotka olivat elossa 31.12.2020.

## Lääkitys

### Lääkinnälliset alaryhmät

Potilaspopulaatio poimimisen lisäksi lääkitystietokantaa voidaan käyttää muodostettaessa hoidollisia alaryhmiä diabeteksen hoidon laaturaportteihin. Toisin kuin populaatiota määrittäessä tällöin tulee huomioida vain viimeisen vuoden aikaiset diabeteslääkeostot (verratuna populaation määrityksessä käytettyä koska tahansa historiassa tehtyä diabeteslääkkeen lääkemääräystä).

Diabetes voidaan luokitella diagnoosin perusteella tyypin 1 ja tyypin 2 diabetekseksi, jotka ovat taudin syntymekanismien, komplikaatioiden esiintyvyyden, diabeetikoiden lukumäärän, käytetyn lääkehoitostrategian ja hoidon porrastuksen osalta hyvin erilaisia tauteja. Lääkeostoihin perustuvia alaryhmiä voidaan käyttää esimerkiksi arvioitaessa diabetesdiagnoosien kattavuutta ja laatua. Esimerkiksi tyypin 1 diabetes on määritelmän mukaan insuliinipuutosdiabetes, joten kaikilla potilailla pitäisi olla käytössä insuliini. Tyypin 2 diabeteksen sisällä on ennusteeltaan ja hoitostrategialtaan hyvin erilaisia alaryhmiä. Vertaamalla lääketietoa tietoon potilaan muusta sairastavuudesta ja laboratoriokokeista, pystytään lisäksi analysoimaan, onko esimerkiksi erittäin korkean riskin potilaat tai insuliinipuutoksesta kärsivät potilaat asianmukaisesti tunnistettu ja onko heillä käytössä hoitosuosituksen mukainen lääkitys.

Alla on kuvattu esimerkkejä raporteista, joita voidaan tuottaa PTA:sta löytyviin diabetesdiagnooseihin, lääkkeistöihin ja laboratoriokokeisiin perustuen laadun arvioinnin ja kehittämisen tueksi. Tuloksia arvioitaessa on huomioitava, että ne edustavat populaatiota riittävän kattavasti vain, jos sekä lääkityskattavuus että diagnoosikattavuus on yli 80 % summapopulaatiosta.

Alla olevassa taulukossa lääkinnällinen alaryhmä määriteltiin sen mukaan, mitä lääkkeitä potilas osti vuonna 2020, tai jos potilas oli kuollut ennen sitä, viimeisenä elinvuotenaan. Aineisto sisältää ne potilaat, jotka oli identifioitu PTA:n ICD-10-diagnoosin E10–E14, eReseptin ATC-luokan A10 tai Kelan etuusrekisterin diabeteslääkityksen alennukseen oikeuttavan koodin perusteella tai – toisin kuin muissa tämän luvun analyyseissä – joilla oli Hilmo ICD-10-diagnoosi E10–E14. Taulukossa sarake "Ei dg" osoittaa ne diabeteslääkettä ostaneet potilaat, joille ei ole kirjattu potilastiedon arkistoon ICD-10 E10-14-koodeja. Tyypin 1 diabeetikoita oli 42590 (7 %), tyypin 2 diabeetikoita 379415 (68 %) ja ilman diagnoosia (24 %) potilaista.

Tyypin 1 diabeetikoista 87 % käytti pelkkää insuliinikorvaushoitoa, ja 8 % käytti insuliiniresistenssiä lieventävää tablettilääkitystä insuliinikorvaushoidon ohella. Potilaista vain 3 %:lle ei löytynyt viimeisimmän seurantavuoden aikana diabeteslääkeostoja, mikä voi selittyä esimerkiksi sairaala- tai laitoshoidolla. Vain 2 %:lla oli käytössä pelkkä tabletti- tai GLP1-lääkitys, mikä viittaa siihen, että diagnoosi tai lääketieto on virheellinen. Tyypin 1 diabeetikoiden diagnoosi vaikutti olevan valtaosalla potilailla linjassa hoitomuodon kanssa.

Tyypin 2 ja muista diabetespotilaista 59 % käytti tabletti- tai GLP1-lääkitystä, 20 % käytti yhdistelmähoitoa ja 6 % käytti pelkkää insuliinihoitoa. Lääkeostoja ei löytynyt 14 %:lla potilaista, mikä voi osittain selittyä diagnosikattavuuden puutteilla.

Niistä potilaista, joilta ei löytynyt diabetesdiagnoosia, 77 %:lta löytyi tuore diabeteslääkeosto, mikä viittaa puuttuvaan diagnoosiin eli diagnosikattavuuden vajeeseen. Siinä joukossa, joilla ei ole merkintää diabetesdiagnoosista PTA:ssa, 23 %:lla (23574) ei ollut myöskään tuoreita diabeteslääkeostoja. He ovat todennäköisesti poimiutuneet populaatioon aiempien lyhytkestoisten diabeteslääkeostojen perusteella.

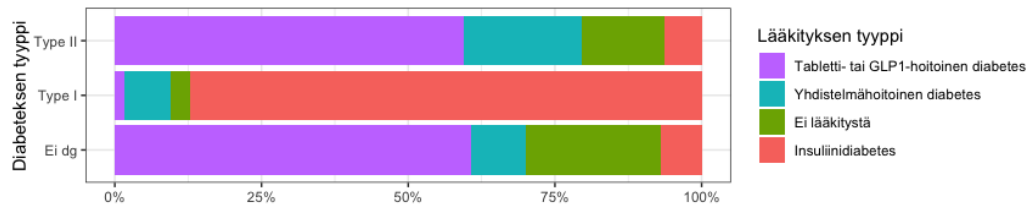
Kuvassa 3.12 on esimerkki diabeteslaaturekisterin raporttiin suunnitellusta kuvasta, jossa kuvan yllä olevan taulukon ydinviesti esitetään tiivistettynä. Kuvasta saa tärkeää tilannekuvaa diabetesdiagnoosien luotettavuudesta ja ajankohtaisista hoitokäytännöistä diabetestyypeittäin.

Kuvassa 3.13 on esimerkki raportista, jossa esitetään HbA1c-tasapainoluokat lääkehoitoryhmittäin tyypin 2 diabeetikoilla. Kuvassa lääkinnällinen ryhmittely perustuu viimeisimmän vuoden aikana tehtyihin lääkeostoihin, ja hoitotasapainoarvio potilaan viimeisimpään HbA1c-laboratoriotulokseen. Hyvässä sokeritasapainossa (HbA1c <53) olevista tyypin 2 diabeetikoista oli suurin osa tablettihoidossa. Sen sijaan huonossa tai erittäin huonossa tasapainossa olevien ryhmässä valtaosa on yhdistelmä- tai insuliinihoidossa, mikä kuvaa insuliinihoidon toteuttamisen haasteita. Kuten myöhemmin raportissa todetaan, tästä aineistosta puuttui vielä niin suuri määrä laboratorioarvoja, että kuvan tuloksiin ei voi luottaa, mutta ne ovat suuntaa antavia ja kertovat siitä, että diabeteksen varhainen diagnostiikka ja hoidon aloitus ovat hyvällä mallilla, ja huonossa hoitotasapainossa olevat tyypin 2 diabeetikot ovat niitä, joille on kehittynyt insuliinipuutos tai joiden insuliinihoidon aloitus on viivästynyt.

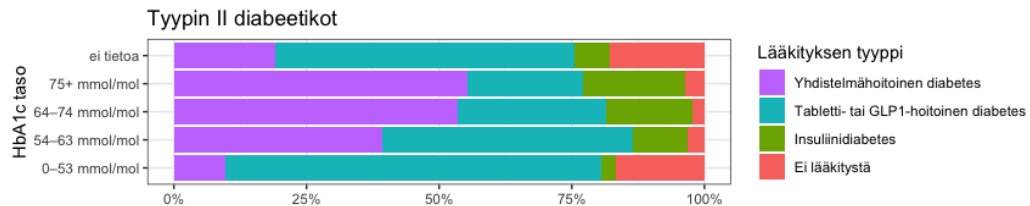
Jatkossa vastaavan kaltaisia kuvia on tarkoitus piirtää myös erikseen esimerkiksi erittäin korkean sydän- ja verisuonisairastavuuden riskipotilailla, jotta pystytään arvioimaan, onko heillä käytössä ennusteeseen vaikuttava diabeteslääkitys, ja onko heidän HbA1c-tasonsa tavoitteissa.



Diabeteksen tyyppi		Type I		Type II		Ei dg		Yhteensä lkm		Yhteensä %	
Lääkinnälliset alaryhmät		lkm	%	lkm	%	lkm	%				
☒ Ei lääkitystä		1444	3,4 %	54012	14,2 %	30574	23,0 %	86030		15,5 %	
☒ Insuliinidiabetes		37092	87,1 %	23639	6,2 %	9101	6,9 %	69832		12,6 %	
– perus- ja ateriainsuliini		31676	74,4 %	14313	3,8 %	4720	3,6 %	50709		9,1 %	
– vain ateriainsuliini		4689	11,0 %	505	0,1 %	545	0,4 %	5739		1,0 %	
– vain perusinsuliini		662	1,6 %	8290	2,2 %	3567	2,7 %	12519		2,3 %	
– vain sekoiteinsuliini		65	0,2 %	531	0,1 %	269	0,2 %	865		0,2 %	
☒ Tabletti- tai GLP1-hoitoinen diabetes		732	1,7 %	225351	59,4 %	80495	60,6 %	306578		55,3 %	
☒ Yhdistelmähoitoinen diabetes		3322	7,8 %	76413	20,1 %	12578	9,5 %	92313		16,6 %	
– ateriainsuliini		287	0,7 %	1060	0,3 %	292	0,2 %	1639		0,3 %	
– monipistoshoido		2622	6,2 %	24178	6,4 %	3191	2,4 %	29991		5,4 %	
– perusinsuliini		400	0,9 %	50509	13,3 %	8918	6,7 %	59827		10,8 %	
– yhdistelmäinsuliini		13	0,0 %	666	0,2 %	177	0,1 %	856		0,2 %	
<b>Kaikki yhteensä</b>		<b>42590</b>	<b>100,0 %</b>	<b>379415</b>	<b>100,0 %</b>	<b>132748</b>	<b>100,0 %</b>	<b>554753</b>		<b>100,0 %</b>	



Kuva 3.12. Diabeteslääkeostoihin perustuvat läikehoitoryhmät suhteessa PTA:n diagnooseihin perustuvaan diabetestyyppiin.



Kuva 3.13 HbA1c-arvoon perustuva diabeteksen hoitotasapaino läikehoitoryhmittäin tyyppin 2 diabeetikoilla.

## Muu lääkitys

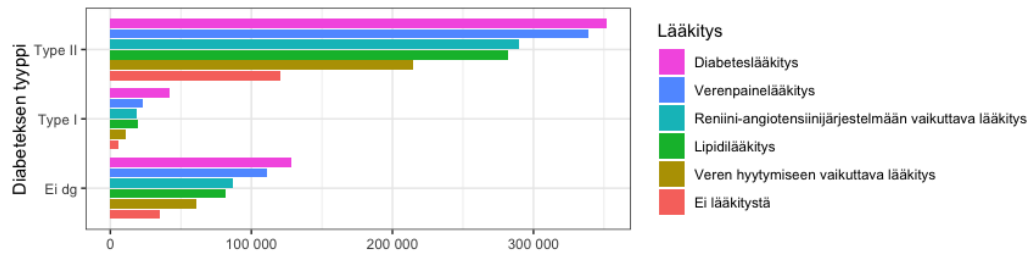
Alla olevassa kuvassa esitetään esimerkki diabetesrekisterin raporttiin suunnitellusta kuvasta muiden ennusteeseen vaikuttavien lääkkeiden käytön prevalenssista diabetestyypeittäin.

Kuva perustuu lääkemääräyksiin.

Tyyppin 2 diabetesta sairastavilla on lisääntynyt sydän- ja verisuonisairauksien riski, ja siksi hoidossa korostuu verenpaineen ja lipidien hoito. Kuvista saa tärkeää tilannetietoa siitä, että tyyppin 2 diabeetikoista valtaosalla on myös verenpaine- ja lipidilääkitys. Jatkossa vastaavan kaltaisia kuvia on tarkoitus piirtää myös erikseen erittäin korkean sydän ja verisuonisairastavuuden riskipotilailla, jotta pystytään arvioimaan, onko heillä käytössä ennusteeseen vaikuttava diabeteslääkitys (ACE-estäjä, SGLT2-estäjä, statiini). Lisäksi voidaan esimerkiksi selvittää, onko erittäin korkean riskin potilaiden LDL-kolesterolitasot tavoitteessa ja onko heillä käytössä lipidilääkitys.

Tyyppin 1 diabeetikoilla korostuu sokeritasapaino ja sen vaikutus retino- ja nefropatian riskiin. Tyyppin 1 diabeetikoilla reniini-angiotensiinijärjestelmään vaikuttavia lääkkeitä ja lipidilääkkeitä tulisi käyttää, jos potilaalle on kehittynyt erittäin korkean riskin merkinä albuminuria. Jatkossa onkin tärkeää arvioida, onko erittäin korkean riskin tyyppin 1 diabeetikoilla käytössä ennustetta parantava lääkitys.

Lääkitysten käytön arviointi suhteessa diabetestyyppiin ja sydänsairauksien riskiin auttaa tunnistamaan myös mahdollisen ylihoidon.



Kuva 3.14 Eri lääkkeitä käyttävien potilaiden määrät diabetesryhmittäin. Verenpainelääkitys kattaa reniini-angiotensiinijärjestelmään vaikuttavan lääkityksen, joka on kuvattu myös omana lääkeryhmänään.

## Hoitomyöntyvyys

Lääkemääräysten ja lääkeostojen avulla on mahdollista karkeasti arvioida myös hoitomyöntyvyyttä. Etenkin tablettimuotoisten diabeteslääkkeiden, verenpaine- ja lipidilääkkeiden osalta tiedetään, että usein huono hoitotasapaino liittyy myös huonoon lääkeadherenssiin eli määrätty lääke jää hankkimatta tai sitä käytetään määrättyä pienemmällä annoksella.

Tässä aineistossa diabeteslääkemääräyksen saaneista potilaista (N=532474) 81 % (N=433626) oli lunastanut lääkettään viimeisten 120 päivän aikana ennen vuoden 2020 loppua tai kuolemaansa. Niistä diabeteslääkemääräyksen saaneista potilaista, jotka olivat vielä elossa vuonna 2021 (N=443497), oli vastaava lunastusprosentti 84 % syys–joulukuulta 2020. Alustavan tiedon mukaan diabeteslääkkeiden komplianssi vaikuttaisi varsin hyvältä. Jatkossa hoitomyöntyvyyttä on tärkeää tarkastella etenkin potilasryhmissä, joissa havaitaan HbA1c- ja LDL-tulosten perustella huono hoitotasapaino. Näin saadaan tärkeää tietoa yleisimmistä syistä hoitotasapainon taustalla, eli johtuuko se todennäköisesti puuttuvasta lääkemääräyksestä, puuttuvasta lääkeadherenssista, liian vähäisestä annoksesta tai vääräntyyppisestä lääkityksestä.

## Sairastavuus ja komplikaatioiden esiintyvyys

Diabetekseen liittyy komplikaatioita, joiden ehkäisy on diabeteksen hoidon keskeinen tavoite. Lisäksi diabeetikoilla saattaa esiintyä muuta sairastavuutta, joka saattaa omalta osaltaan lisätä komplikaatioiden riskiä tai vaikeuttaa komplikaatioiden ehkäisyä, esim. psyykkiset sairaudet tai uniapnea. Tämän vuoksi diabetesrekisterissä raportoidaan komplikaatioiden ja muiden sairauksien esiintyvyyttä ja ilmaantuvuutta. Tietoja komplikaatioiden esiintyvyydestä on saatavilla kansainvälisistä ja kansallisista epidemiologisista tutkimuksista, ja niiden perusteella voidaan karkeasti arvioida datan ja analyysien luotettavuutta.

Diabeteksen komplikaatioiden ja diabeteksen hoitoon vaikuttavien sairauksien esiintyvyyttä arvioidaan luokittelemalla potilaalta löytyvät pää- ja sivudiagnoosit sairausryhmittäin. Mikäli potilaalla on tarkasteluajan loppuun mennessä esiintynyt jokin sairausryhmään luokitteleva diagnoosi tai toimenpide, potilaan katsotaan saaneen ko. komplikaation.

## Sydän- ja verisuonisairaudet

Sydän- ja verisuonisairaudet eli kardiovaskulaarisairaudet (KVS) valittiin validoitavaksi sairausryhmäksi, koska se on yksi tärkein diabeteksen ehkäistävissä oleva komplikaatio. Sydän- ja verisuonisairauksien esiintyvyys (KVS-prevalenssi) vaikuttaa myös LDL-tavoitteisiin. Sairausryhmän määrittelyyn käytetään sekä diagnooseja että toimenpiteitä. Joitakin toimenpiteitä

tehdään myös radiologien toimesta, jolloin toimenpiteet kirjautuvat kuvantamisen järjestelmiin. Tämän vuoksi haluttiin selvittää, poimiutuvatko toimenpideradiologiset toimenpiteet PTA:oon.

Diagnosi- ja toimepidekoodien löytymistä verrattiin PTA- ja Hilmo-aineistojen välillä. Todettiin, että valtaosa KVS-luokitteluun vaadittavista diagnooseista ja toimepidekoodista löytyy kattavasti PTA-aineistosta. Tehtyjen havaintojen perusteella tarkennettiin KVS-prevalenssin arviointiin tarvittavia diagnooseja ja toimenpiteitä sekä arvioitiin niiden esiintymisen yleisyyttä. Sairastavuutta koskevat havainnot on kuvattu tarkemmin DM-sairastavuus-Excelin Sairastavuus-välilehdellä.

Alla on kuvattu KVS-prevalenssissa käytetty määrittely ja siihen tehdyt korjaukset:

- Sepelvaltimotauti tai sydämen vajaatoiminta: ICD-10: I20-25, I46, I50, tai FNA, FNC, FNE, FN2 alkuiset toimenpiteet. Nämä tmp-koodit löytyivät pääsääntöisesti PTA:sta hiukan kattavammin kuin Hilmosta
- Aivoverenkierron sairaudet: ICD-10: I63 - I66, (paitsi I63.6) ja G45, tai PA ja PB-alkuiset toimenpiteet. Havaittiin alkuperäisessä määrittelyssä esiintynyt määrittelyvirhe, joka korjattiin. Tähän luokkaan otetaan mukaan vain valtimoiden avaamiseen tarkoitetut toimenpiteet **PFA14**, **PA2VT** ja **PAN14** (ensimmäinen eli a. carotiksen trombendarerektomia on yleisin, muut ovat tukoksen mekaanista poistoa tarkoittavia). Muut alkuperäisellä määrittelyllä löytyvät toimenpiteet ovat pääosin radiologisia tutkimuksia.
- Jalkojen verenkierron sairaudet: ICD-10: I70.2, I79.2, E10.5, E11.5, E12.5, E13.5, E14.5 tai PD, PE, PF, PG alkuiset toimenpiteet. Alkuperäisessä määrittelyssä havaittiin virhe, ja todettiin, että toimenpiteistä mukaan otetaan vain valtimoiden avaamistoimenpiteet PE1AT (PTA/Hilmo 50 %), PF1AT (PTA/Hilmo 50 %), **PEF23 (PTA/Hilmo >100 %)**, PE1YT (PTA/Hilmo 50 %), PD3YT (PTA/Hilmo 70 %), PF1BT (PTA/Hilmo 50 %), PE1BT (PTA/Hilmo 50 %), **PEN23 (PTA/Hilmo >100 %)**, PD3AT (PTA/Hilmo 60 %), PEH57, PEF25, PEF24, PEH56, PDQ20, PEN24, PG2YT, PFH66. Tässä ovat mukana vain ne toimenpiteet, joiden esiintyvyys on yli 50. Havaittiin, että valtimotaudin hoitotoimenpiteitä tehdään osittain radiologisina toimenpiteinä ja kirjautuvat radiologian ns. RIS-tietojärjestelmiin. Näistä järjestelmistä toimepidekoodit tallentuvat PTA:oon, jos järjestelmä on sertifioitu A-luokan järjestelmäksi.

Seuraava taulukko kuvaa em. sairauksien esiintyvyyttä PTA:ssa diabetesdiagnoosin saaneiden keskuudessa. Mukana kaikki E10–E14-koodilla PTA:han kirjatut potilaat, myös sittemmin kuolleet. Jatkossa vastaava taulukko tehdään siten, että jakajana käytetään summapopulaatiota ja tulokset esitetään hyvinvointialueittain sekä kaikki KVS-sairaudet yhdistäen. On hyvä huomata, että diabetesdiagnoosien kattavuudessa havaittu ongelma koskee todennäköisesti myös sairastavuuden arviointiin käytettäviä diagnoosikodeja, ja siksi matalan diagnoosikattavuuden alueella KVS-prevalenssi voi olla todellisuutta matalampi.

KVS-prevalenssia käytetään kardiovaskulaarisairauksien esiintyvyyden arvioinnin lisäksi hoitotavoitteiden raportoinnissa. Esimerkiksi LDL-kolesterolitasot näyetään erikseen erittäin korkean riskin diabeetikoilla, jolloin ryhmään kuuluvat ne potilaat, joilla on KVS-sairaus, valkuaisvirtsaisuus tai eGFR alle 30.

Diagnoosien ja toimenpiteiden päivämäärien avulla voidaan arvioida viimeisen kahden vuoden aikaista insidenssiä (viimeisen kahden vuoden aikana ko. diagnoosilla esiintyneitä hoitajaksoja), mikä on komplikaatioiden ehkäisyn onnistumista kuvaava tärkeä laadun indikaattori.

DM tyyppi	Sairaus	Lkm	Potilaita	Prevalenssi
DM1	Sepelvaltimotauti, sydämen vajaatoiminta	4628	42590	10,9 %
DM1	Aivoverenkierron sairaudet	504	42590	1,2 %
DM1	Jalkojen verenkierron sairaudet	2416	42590	5,7 %
DM2	Sepelvaltimotauti, sydämen vajaatoiminta	107282	379415	28,3 %
DM2	Aivoverenkierron sairaudet	5161	379415	1,4 %
DM2	Jalkojen verenkierron sairaudet	26867	379415	7,1 %

## Retinopatia eli verkkokalvosairaus

Proliferatiivisen retinopatian (verkkokalvosairaus, jolle ovat ominaisia uudissuonikasvu ja arkipudoksen muodostuminen) hoito löytyi valtaosin toimenpidekoodeilla CKC12 tai CKD05. Kokeiltiin myös laajempaa hakua: CKC\* ja CKD\*, mikä tuotti suuremman variaation erilaisia retinopatian hoitotoimenpiteitä, mutta ei mitään asiaan täysin kuulumatonta. Valtaosassa toimenpiteitä PTA:sta löytyneiden toimenpidekoodien esiintyvyys oli suurempi kuin Hil-mossa todennäköisesti siksi, että PTA:n tietoihin tuli mukaan myös yksityissektorin kirjaamat toimenpidekoodit. Jatkossa retinopatia määritellään seuraavasti: mikäli potilaalla on joskus esiintynyt CKC\* tai CKD\*-toimenpide tai diagnoosi H36 (pois lukien H36.8), H28.0, H43.1, H45.0, H54, H40.5, H42.0, E10.3, E11.3, E12.3, E13.3, tai E14.3, hänellä on retinopatia (retinopatian esiintyvyys). Toimenpiteiden CKC\* tai CKD\* päivämäärien avulla voidaan arvioida viimeisen kahden vuoden aikaista akuuttihoidon tarpeen insidenssiä, mikä on toinen tärkeä laadun indikaattori.

Prosessimittarina ja resurssien käytön mittarina suunniteltiin poimittavaksi silmänpohjakuvaus ja sen ajankohta (toimenpidekoodit CL2S4 tai CL2L4). Tarkoitus oli arvioida, toteutuuko silmänpohjakuvaus hoitosuosituksen mukaisesti vähintään 2 vuoden välein. Näistä CL2S4 löytyi 75 %:sti PTA:sta (n=2386 Hil-mossa ja n= 1786 PTA:ssa). Kumpikin toimenpidekoodi esiintyi hyvin harvalla potilaalla suhteessa potilaiden määrään eli todennäköisesti valtaosa kuvantamisesta jää kirjautumatta. Näin ollen PTA:sta löytyvän silmänpohjakuvaustiedon pohjalta ei voi tehdä johtopäätöksiä silmänpohjakuvausten toteutumisesta. Tieto silmänpohjakuvausten toteutumisesta Käypä hoitosuosituksen mukaan ja toisaalta oikeasta kohdentumisesta korkean riskin potilaisiin olisi jatkossa tärkeä laatuindikaattori.

## Diabeteksen jalkakomplikaatiot

Jalkahaavan hoitoon tarvittavia toimenpiteitä selvitettiin nilkan, jalkaterän ja alaraajan haavan revisiotoimenpiteillä NHW\*, NHK99, NHS99, NHS20, QDB05, QDG20, QDG30, QDG99. Näistä valtaosa oli harvoin käytettyjä. Vain QDB05 ja QDG20 eli haavarevisiot esiintyivät niin yleisesti, että ne kannattaa ottaa jatkossa mukaan osoittamaan viimeisen kahden vuoden aikana tehtyjen kirurgisten revisioiden insidenssiä. Havaittiin näiden toimenpiteiden hyvät kattavuus PTA:ssa, ja se oli jopa hiukan parempi kuin Hil-mossa. Eri tasoiset amputaatiot ovat toinen tärkeä jalkahaavojen akutisoitumisen insidenssiä kuvaava muuttuja: Nilkan yläpuolinen amputaatio NFQ10, NFQ20, NGQ10, NGQ20, NHQ10, NHQ20, NFQ48 tai NGQ48 ja nilkan alapuolinen amputaatio NHQ30, NHQ40, NHQ60. Amputaatioiden toimenpidekoodeissa havaittiin hyvä tietojen kattavuus PTA:ssa.

Seuraavassa taulukossa esitetään retinopatian ja jalkakomplikaatioiden esiintyvyys DM1- ja DM2-potilailla:

DM tyyppi	Toimenpide	Lkm	Potilaita	Esiintyvyys
DM1	Silmäsairaudet ja retinopatian hoito	18442	42590	43,3 %
	Nilkan, jalkaterän ja alaraajan haavan revisiot	888	42590	2,1 %
	Nilkan yläpuolinen amputaatio	327	42590	0,8 %
	Nilkan alapuolinen amputaatio	748	42590	1,8 %
DM2	Silmäsairaudet ja retinopatian hoito	34661	379415	9,1 %
	Nilkan, jalkaterän ja alaraajan haavan revisiot	5437	379415	1,4 %
	Nilkan yläpuolinen amputaatio	2092	379415	0,6 %
	Nilkan alapuolinen amputaatio	3772	379415	1,0 %

## Munuaisten vajaatoiminta

Diabeteslaaturekisteripilottiaineistosta tehdyn selvityksen perusteella diabeteksen munuais-taudin esiintymisessä on luotettavampaa käyttää laboratorioaineistoon pohjautuvaa eGR-tietoa kuin diagnoosikirjauksia, koska diagnoosien kirjauskattavuus on hyvin matala. Poikkeus on loppuvaiheen munuaisten vajaatoiminta (ESRD), joka tulisi määrittää seuraavasti: Hemodialyysi (Z49.1), peritonealidialyysi (Z49.2), munuaissiirre (Z94.0) tai viimeinen eGFR alle 15 ml/min/1.73m<sup>2</sup>

Alla on kuvattu eGFR-tietoon perustuva munuaisten vajaatoiminnan esiintyminen (ryhmä 3 tai ryhmä 4) diabetestyypeittäin. Tässä vaiheessa laboratoriotietoja puuttuu vielä PTA:sta, joten luokittelun tuloksiin tulee suhtautua varauksella.

ESRD on kallein diabeteksen komplikaatio ja sen esiintyvyydestä ja ilmaantuvuudesta on tärkeää olla hyvinvointialuekohtaista tietoa. Lisäksi eGFR-ryhmä 15–59 on merkittävä ryhmä sikäli, että hyvällä sokeritasapainolla ja ennustetta parantavalla lääkityksellä (ACE-estäjä, SGLT2-estäjä) taudin etenemistä ESRD:ksi pystytään ehkäisemään. Jatkossa hoitotuloksia on hyvä esittää myös eGFR-mukaisissa ryhmissä.

Tyyppi	N	%
<b>DM1</b>	<b>42590</b>	<b>10,1 %</b>
1: eGFR ei tiedossa	10932	25,7 %
2: eGFR < 15 tai dialyysihoito	1521	3,6 %
3: eGFR 15–59	3805	8,9 %
4: eGFR 60+	26332	61,8 %
<b>DM2</b>	<b>379415</b>	<b>89,9 %</b>
1: eGFR ei tiedossa	99218	26,2 %
2: eGFR < 15 tai dialyysihoito	9912	2,6 %
3: eGFR 15–59	96836	25,5 %
4: eGFR 60+	173449	45,7 %
<b>Kaikki yhteensä</b>	<b>422005</b>	<b>100,0 %</b>

## Suusairaudet

Suusairauksia kuvaavat diagnoosi- ja toimenpidekoodit eivät löydy Terveystietokannasta. Suusairaudet löytyvät PTA:sta diagnoosilla: K04 (periapikaalinen, juuri) tai toimenpidekoodilla SGA, K05 tai SDA01–SDA02 (lievä), SDA03–SDA14 (muu parodontiitti) sekä T84 (peri-implanttisairaudet).

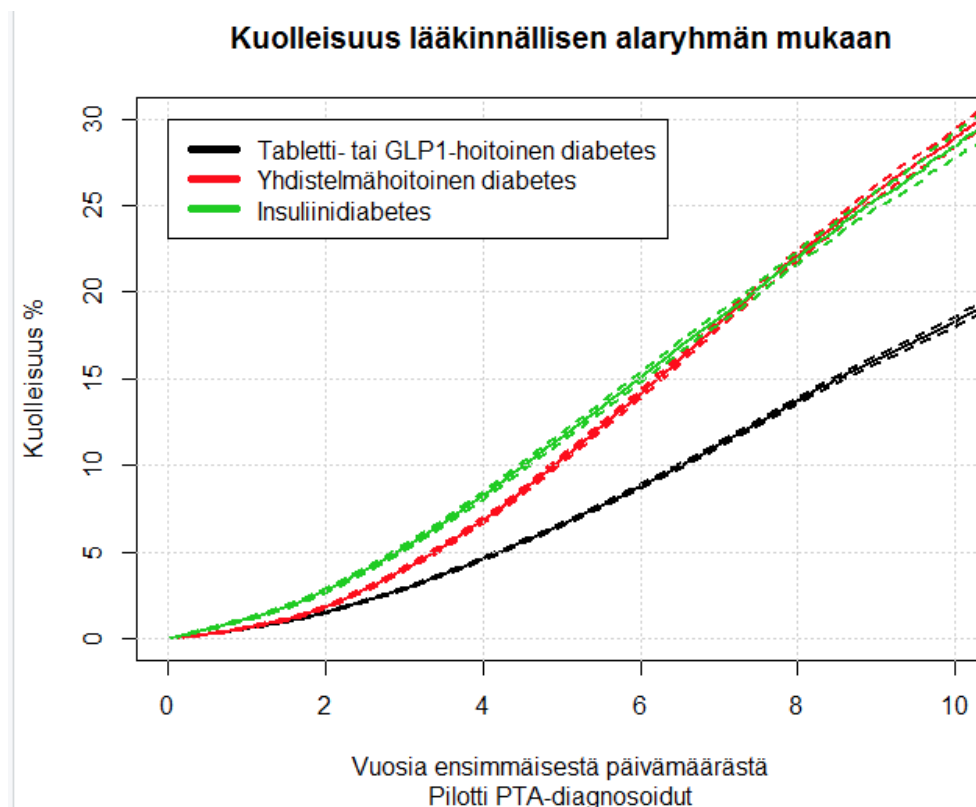
## Lihavuus

Lihavuuden hoitoon käytetyt bariatrisen eli lihavuutta hoitavan kirurgian JDF-toimenpidekoodit löytyvät kattavasti PTA:sta.

## Kuolleisuus

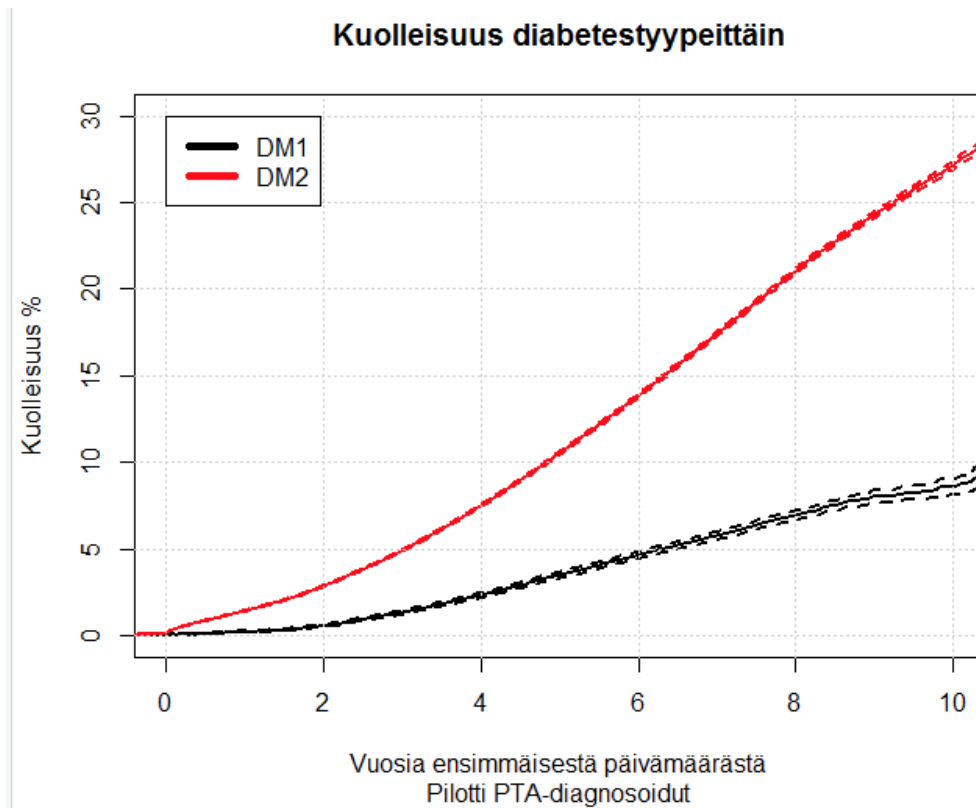
Diabeteksen hoidon keskeinen tavoite on diabetekseen liittyvän ylikuolleisuuden välttäminen. Alla esitetään esimerkkeinä raportteihin suunniteltuja kuvia kuolleisuudesta. Seurannan alkuhetki ("diagnoosi") on varhaisin päivämäärä, jolloin potilaalta on havainto diabeteksestä jossain käytetyistä tietolähteistä (PTA, lääkemääräys, etuus). Jos PTA-diagnoosiin on kirjattu tätäkin varhaisempi diabeteksen toteamispäivä, niin silloin se toimii alkuhetkenä. Kuvien tuloksiin tulee suhtautua varauksella, koska tosiasiallisesti diabeteksen kesto on usein pidempi kuin näissä analyyseissä käytetty.

Kuvassa 3.15 nähdään, että insuliini- ja yhdistelmähoitoisilla potilailla on selvästi huonompi ennuste kuin tablettihoitoisilla diabeetikoilla. Kuvasta 3.16 nähdään että tyypin 2 diabeetikoilla on korkeampi kuolleisuus kuin tyypin 1 diabeetikoilla, ja kuvasta 3.17 nähdään, että kardiovaskulaarisairaus lisää huomattavasti kuolleisuutta. Kuvasta 3.19 nähdään, että tyypin 2 potilaiden kuolleisuus vaihtelee hyvinvointialueittain.



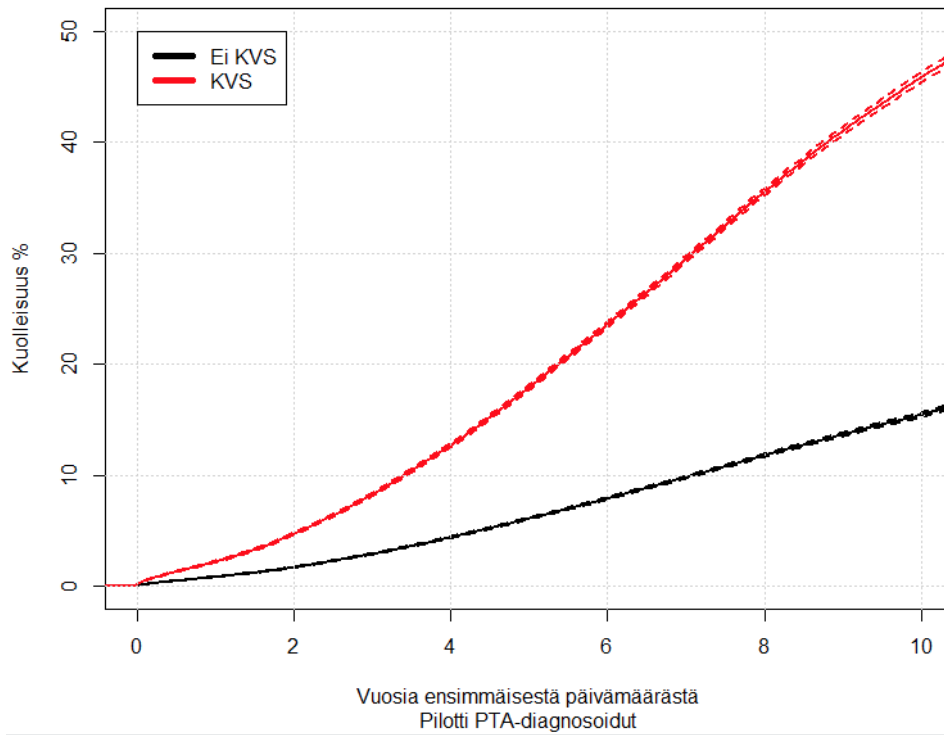
Kuva 3.15 PTA-diagnosoitujen potilaiden kuolleisuus 10 vuoden aikana ensimmäisestä diabetesmerkinnästä lääkinällisen pääryhmän mukaan jaoteltuna (Kaplan–Meier-estimaatit). Kuviin on piirretty myös 95 %:n luottamusväli.

Vastaava kuva diabeteksen tyypin suhteen:

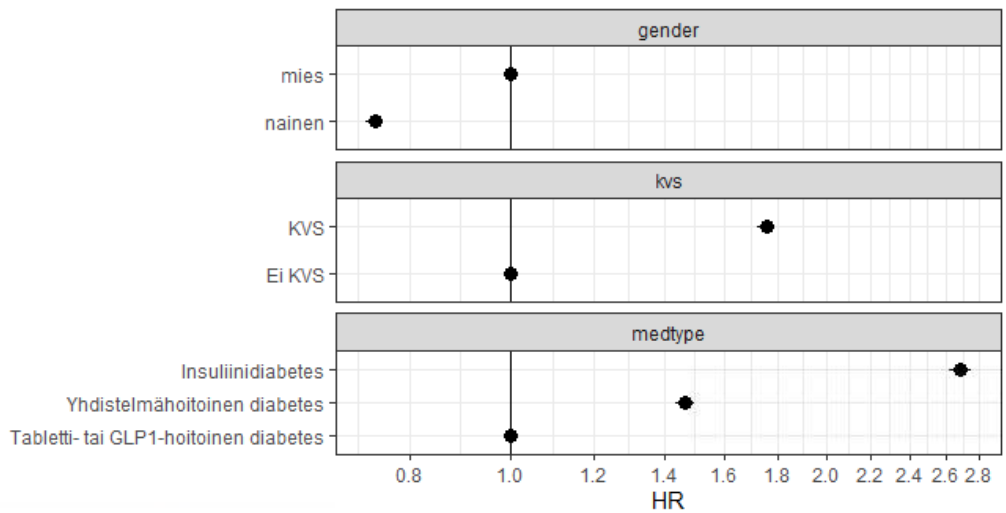


Kuva 3.16 PTA-diagnosoitujen potilaiden kuolleisuus 10 vuoden aikana ensimmäisestä diagnosoimarkinnästä diabetestyypeittäin jaoteltuna (Kaplan–Meier-estimaatit). Kuviin on piirretty myös 95 %:n luottamusväli.

## Kardiovaskulaarisairastavuus ja kuolleisuus

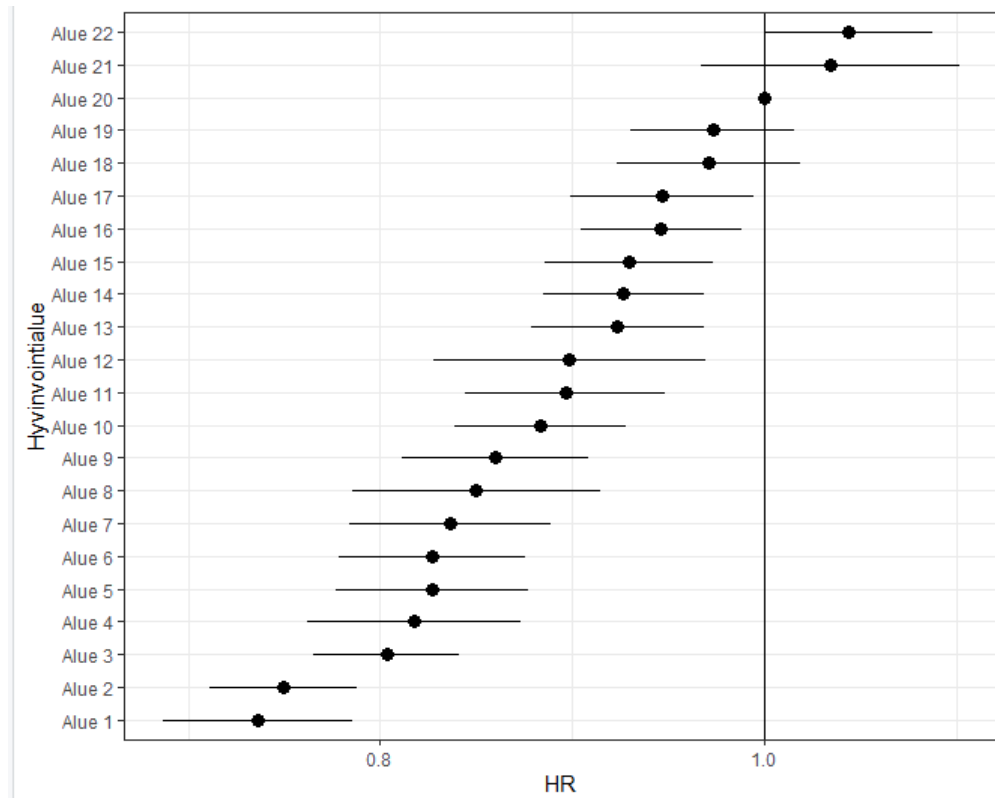


Kuva 3.17 Kuolleisuus potilailla joilla on kardiovaskulaarisairaus ja niillä, joilla ei ole kardiovaskulaarisairautta.



Kuva 3.18 Coxin elinaikamallin ikävakioiduista kuolleisuuden riskitiheysuhteista ("hazard ratios", HR) naisille suhteessa miehiin, kardiovaskulaarisairautta poteville (KVS) sekä eri lääkinnällisille alaryhmille. Aineistona on tyypin 2 diabetesta sairastavat, PTA-diagnosoidut, Manner-Suomessa asuvat potilaat. Kuvissa myös piste-estimaatin lisäksi 95 %:n luottamusväli.





Kuva 3.19 Tyypin 2 potilaiden kuolleisuus hyvinvointialueittain. Hyvinvointialueiden vakioidut riskitiheysuhteet ("hazard ratios", HR) on estimoitu Coxin elinaikamallilla, jossa ikä, sukupuoli, kardiovaskulaarisairastavuus sekä lääkinälliset alaryhmät on vakioitu sekoittavina tekijöinä. Kuvaan on piirretty piste-estimaatin lisäksi myös tämän 95 %:n luottamusväli. Hyvinvointialueiden nimet on pseudonymisoitu, koska aineistoon liittyy vielä valikoituvuutta eräiden alueiden alhaisemman diagnoosikattavuuden vuoksi. Lisäksi elinajan seurannan alun määrittämisessä saattaa edelleen alueellisesti esiintyä systemaattisia eroja. Sekoittavien tekijöiden lista ei myöskään ole kattava.

## Laboratoriokokeet

### Laboratoriokokeiden tulosten kattavuudet

Diabeteksen hoidon keskeisiä laatuindikaattoreita ovat laboratoriotutkimuksista HbA1c (sokeritasapaino), LDL-kolesteroli, U-AlbKre (valkuaisvirtsaus) ja eGFR (munuaisten toiminta, joka lasketaan P-Krea arvon pohjalta). Laboratoriokokeet ovat tärkeitä laatuindikaattoreita, koska niihin vaikuttamalla voidaan vähentää diabeteksen komplisoitumisen riskiä, ja ne reagoivat nopeasti hoidossa tapahtuviin muutoksiin. Diabetesrekisterin kannalta kriittisintä on, saammeko poimittua diabeteksen hoidon kannalta keskeiset laboratoriokokeet PTA:sta luotettavasti ja kattavasti.

Kunkin diabeteksen seurannassa käytetyn laboratoriokokeen osalta arvioimme ensin, millä osuudella PTA:sta diagnoosin perusteella löytyneistä diabeetikoista löytyi diabeteksen seurannassa käytetty laboratoriokoe vuosilta 2019–2020.

Lisäksi tutkimme laboratoriokokeiden löytymisen kattavuutta vertaamalla Kanta PTA-tietoaan laboratorioaineistoa pilottivaiheessa kolmesta sairaanhoitopiiristä kerättyyn laboratorioaineistoon. Pilottivaiheen sairaanhoitopiirit ja laboratoriokokeiden tietolähteet olivat Pohjois-Karjala (Siun soten potilastietojärjestelmä), Pirkanmaa (Fimlabin WebFimlab) ja Keski-

Suomi (Fimlabin WebFimlab sekä Saarikasta ja Wiitaunioinista CGI:n tuottama tiedosto). Kumpaankin aineistoon otettiin mukaan ne potilaat, joiden kotikunta oli näillä hyvinvointialueilla vertailujaksolla. Vertailujakso oli Pohjois-Karjalassa 17.9.2018-16.9.2020, Pirkanmaalla ja Keski-Suomessa 1.9.2018-31.8.2020. Jos tutkimustulosarvoa ei löytnyt, tutkimus hylättiin aineistosta (tyypillisesti peruttuja tai epäonnistuneita laboratoriokokeita). PTA-aineiston potilaat olivat niitä, joille löytyi PTA:sta E10-E14 diagnoosi ja jotka löytyivät VTJ:stä. PTA-aineistoon laboratoriokokeet oli poimittu vain Kuntaliiton koodilla 6128 (HbA1c), 4599 (fP-LDL), 4600 (P-Krea) ja 4511 (U-AlbKre). Pilottiaineistoon potilaat oli poimittu Hilmosta ja AvoHilmosta löytyneiden vastaavien diagnoosinumeroiden perusteella. Pilottiaineistoon poimittiin laboratoriokokeisiin muitakin kuin Kuntaliiton koodilla löytyviä laboratoriotutkimuksia: LDL-koodin 4599 lisäksi poimittiin P-SLDL 532, P-SLDL114 ja P-LDLlask 9535; Krea-koodin 4600 lisäksi poimittiin 2142, s-krea 2143 ja krea+GFR 6002, 6354; U-AlbKre-koodin 4511 lisäksi poimittiin 9522). Näin ollen vertailtavat populaatiot ja poimitut laboratorioaineistot eivät ole määrittelyiltään täysin identtiset, joten aineistojen ei odotettu olevan täysin identtiset. Vertailulla pyrittiin kuitenkin tunnistamaan selkeät erot laboratoriokoekestävyydestä ja niiden syyt.

Tilannetta tutkittiin tarkemmin kuntakohtaisesti pilottirekisterissä mukana olleilta hyvinvointialueilta. Pilottirekisteriin osallistuneet sairaanhoitopiirit vastaavat pääosin tulevia hyvinvointialueita kahta poikkeusta lukuun ottamatta. Jämsä on Keski-Suomen hyvinvointialueen kunta, vaikka se pilotissa kuului Pirkanmaan sairaanhoitopiiriin. Pirkanmaan hyvinvointialueeseen liittyy Punkalaidun, joka ei kuulu Pirkanmaan sairaanhoitopiiriin, minkä takia se ei näy tutkittujen kuntien joukossa.

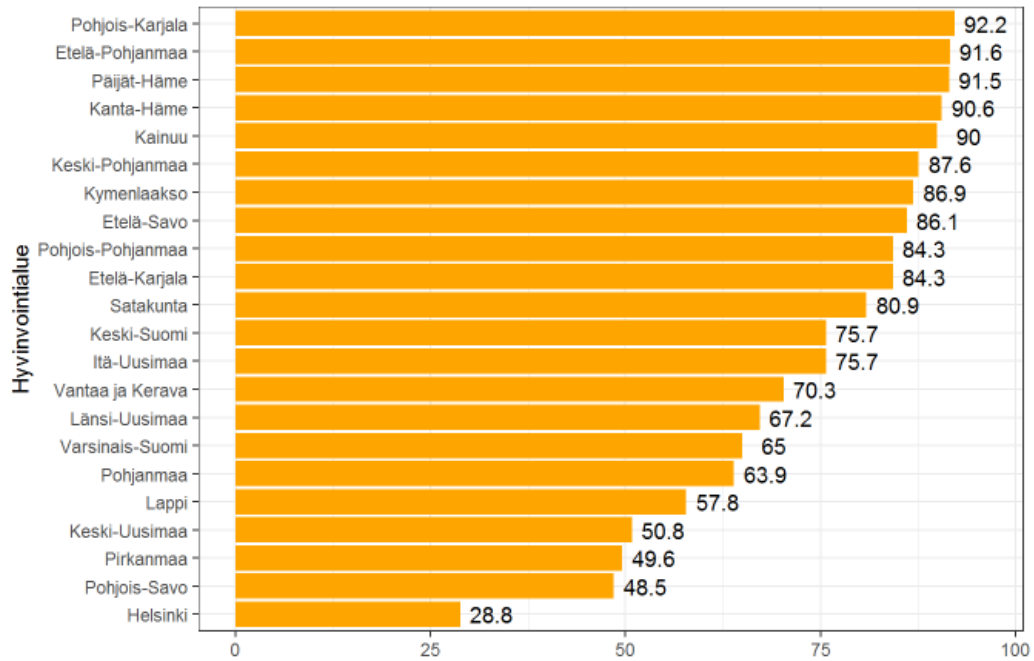
#### **PTA-aineistosta poimitut HbA1c-kokeet (6128)**

Kuvassa 3.20 esitetään PTA-aineistosta löytyneiden HbA1c-kokeissa käyneiden potilaiden määrä 100 PTA-aineistosta diagnoosin perusteella löytynyttä diabeetikkoa kohti hyvinvointialueittain. Pilottiaineiston perusteella tiedetään, että HbA1c kokeessa käy yli 75 % potilaista, joten tulosten perusteella heräsi epäily siitä, että PTA:sta ei löydy kaikkien hyvinvointialueiden kaikkia tutkittuja HbA1c-arvoja. Näin ollen PTA:sta poimittujen HbA1c-arvojen tietoihin ei voinut luottaa kaikilla Hyvinvointialueilla.

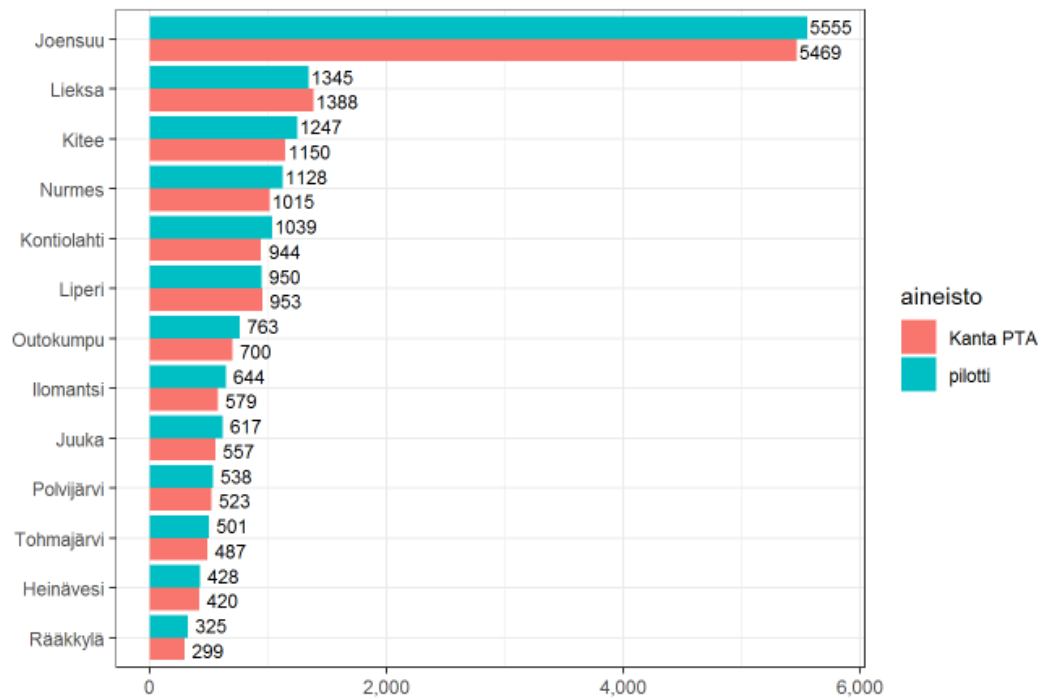
Pilottirekisterissä mukana olleista Hyvinvointialueista Pohjois-Karjalasta (Siun sote) HbA1c arvot näyttivät löytyvän erinomaisen kattavasti (kuva 3.21), kun taas Pirkanmaan alueella HbA1c mittauksia löytyi vain 52 %:lta potilaista (kuva 3.21) ja Keski-Suomessa 73 %:lta (kuva 3.22). Pilottirekisterin ja PTA-aineiston laboratoriopoinnoissa oli kummassakin käytetty vain Kuntaliiton koodia 6128.

Pirkanmaalla oli huomattavaa vajetta HbA1c-kokeiden löytymisessä PTA:sta, kun tulosta verrattiin pilottiaineistossa suoraan WebFimlabista löytyneisiin laboratoriokokeisiin (kuva 3.22). Lisäksi kuntien välillä oli huomattavaa vaihtelua. Esimerkiksi Nokialla, Virroilla, Ikaalisissa ja Hämeenkyrössä HbA1c arvot löytyvät kattavasti PTA:sta. Sen sijaan valtaosassa kunnista HbA1c arvot eivät kattavasti löytyneet PTA:sta.

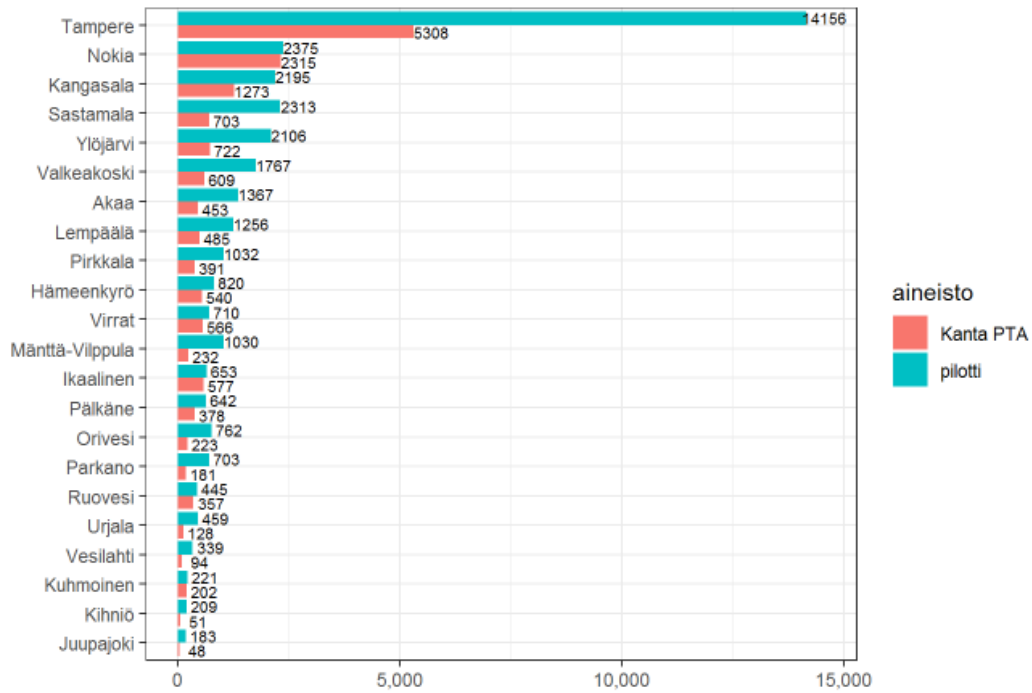
Keski-Suomessa oli samankaltainen tilanne kuin Pirkanmaalla eli Jämsää ja Äänekoskea lukuun ottamatta PTA:sta puuttui HbA1c-kokeita (Kuva 3.23).



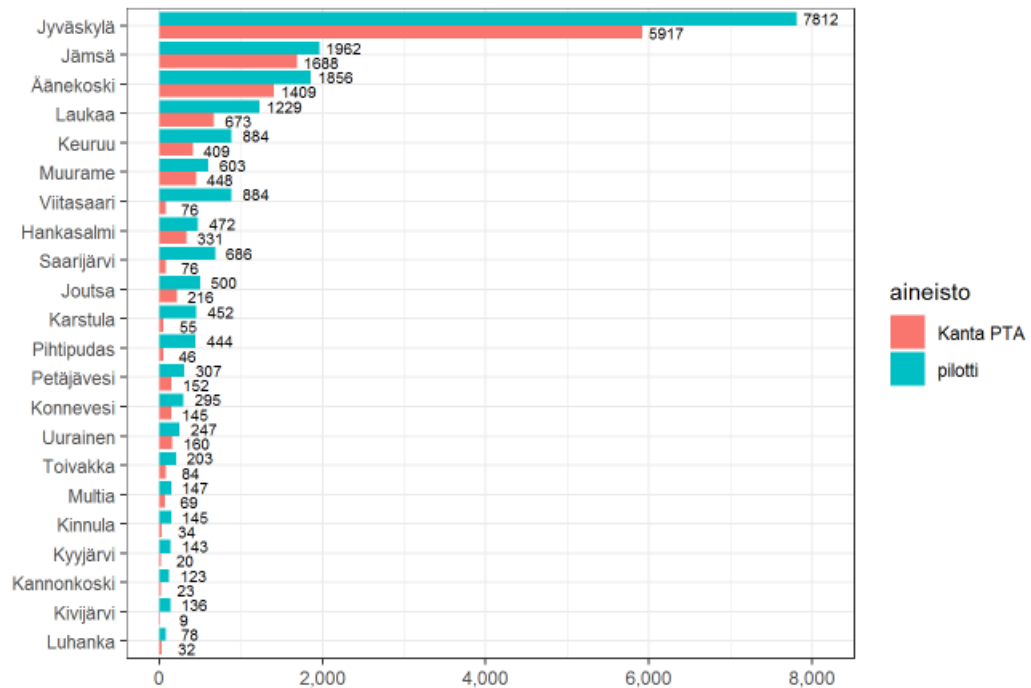
Kuva 3.20: Niiden potilaiden osuus PTA-aineiston potilaista, joille löytyi HbA1c-laboratoriotulos (%)



Kuva 3.21: Pohjois-Karjalan hyvinvointialueella kotikunnittain niiden henkilöiden lukumäärä, joille löytyy HbA1c- laboratoriotulos



Kuva 3.22: Pirkanmaan hyvinvointialueella kotikunnittain niiden henkilöiden lukumäärä, joille löytyy HbA1c-kokeen laboratoriotulos

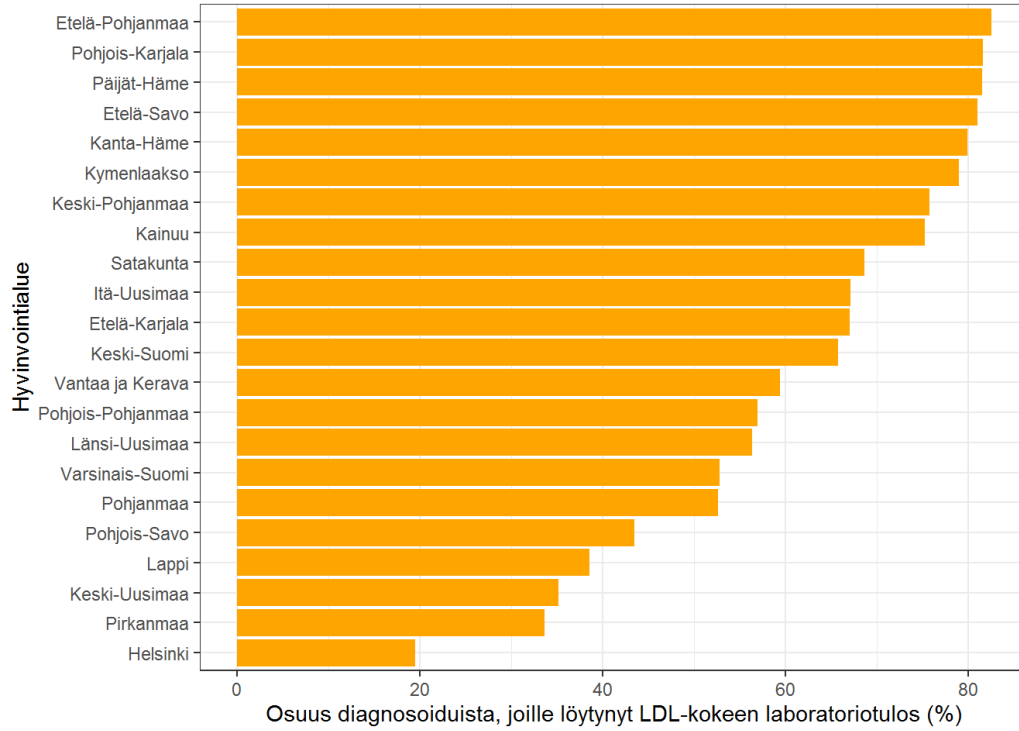


Kuva 3.23: Keski-Suomen hyvinvointialueella kotikunnittain niiden henkilöiden lukumäärä, joille löytyy HbA1c-laboratoriotulos

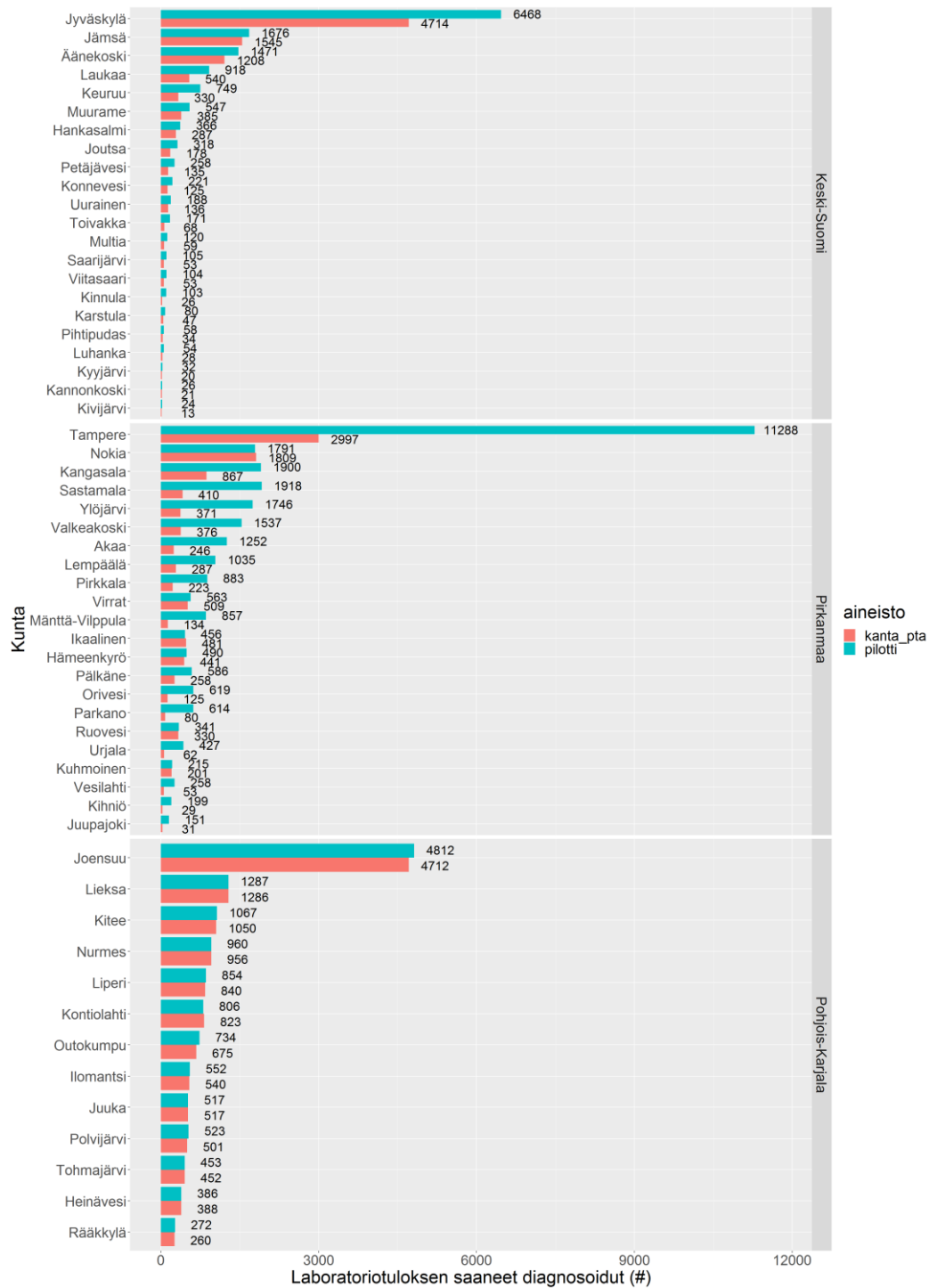
### PTA-aineistosta poimitut fP-LDL-kokeet (4599)

Kuvassa 3.24 esitetään PTA-aineistosta löytyneiden LDL-kolesterolikokeissa käyneiden potilaiden määrä 100 PTA-aineistosta diagnoosin perusteella löytynyttä diabeetikkoa kohti hyvinvointialueittain. Myös LDL-kolesterolissa esiintyi matalaa kattavuutta pääosin samojen sairaanhoitopiirien alueilla kuin HbA1c-kokeessakin. Kuvassa 3.25 verrataan pilottiaineistosta ja PTA-aineistosta löytyviä laboratoriotuloksia. Löydökset ovat pääosin samat kuin HbA1c:n

osalta eli Pohjois-Karjalan alueella laboratorioarvot löytyvät kattavasti kaikista kunnista, kun taas Keski-Suomessa ja Pirkanmaalla LDL-kolesteroliarvot löytyvät vain tiettyjen kuntien alueilta.



Kuva 3.24 Niiden potilaiden osuus PTA-aineiston potilaista, joille löytyi LDL-kolesterolitulos (%)



Kuva 3.25 Keski-Suomen, Pirkanmaan ja Pohjois-Karjalan hyvinvointialueilla kotikunnittain niiden henkilöiden lukumäärä, joille löytyy LDL-kolesterolitulos.

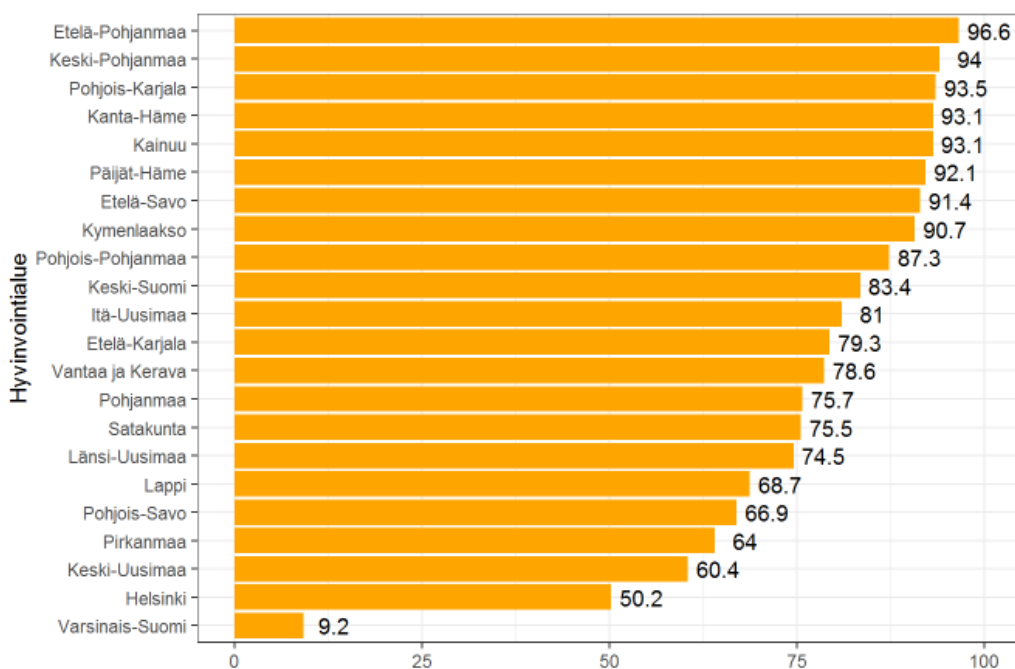
### PTA-aineistosta poimitut P-Krea-kokeet (4600) ja U-Alb/Kre-kokeet (4511)

Kuvassa 3.26 esitetään P-Krea-kokeessa ja kuvassa 3.27 U-Alb/Kre-kokeessa käyneiden potilaiden lukumäärä suhteessa 100 PTA:sta löytyneseen diabeetikkoon hyvinvointialueittain. Mittauskattavuuden vajeus on P-krea-kokeessa havaittavissa samoilla hyvinvointialueilla kuin HbA1c-kokeessa. U-Alb/Kre-kokeessa havaittiin kaikissa sairaanhoitopiireissä huomattavasti alhaisempi mittauskattavuus kuin muissa diabeteksen seurannassa käytetyissä laboratoriotuloksissa, mikä voi johtua osittain hoidollisista ja osittain teknisistä syistä.

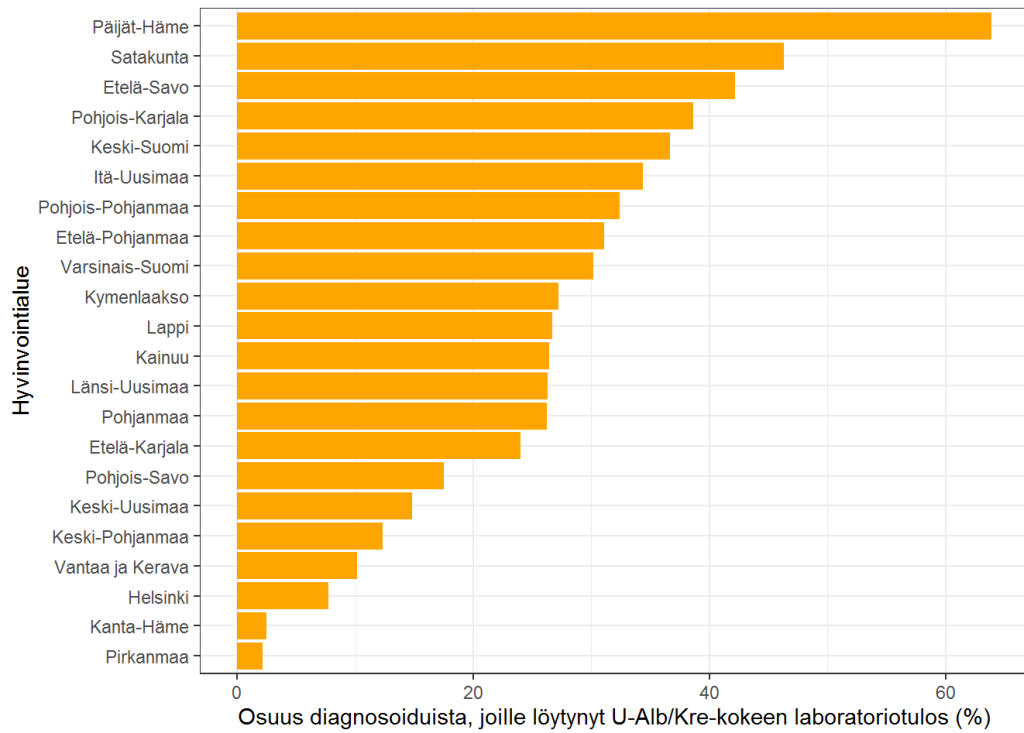
Kuvassa 3.28 verrataan pilottiaineistosta ja PTA-aineistosta löytyviä U-Alb/Kre-laboratoriokokeita. U-Alb/Kre-kokeessa erot PTA-aineiston ja pilottiaineiston kattavuudessa olivat vielä

suuremmat kuin verikokeiden kattavuudessa. U-AlbKre koetta ei löytynyt PTA-aineistoista niidenkään kuntien alueilta, joista löytyivät muut tutkitut diabeteksen seurannassa käytetyt laboratorioskokeet. Esimerkiksi Keski-Suomen alueella Jämsästä ja Pirkanmaalla Nokialta ei löytynyt juuri lainkaan U-AlbKre-tuloksia, vaikka verikokeiden kattavuus oli varsin hyvä. Keski-Suomen ja Pirkanmaan kunnista ainoastaan Äänekosken alueella UalbKre-arvoja näytti löytyvän lähes samalta määrältä potilaita kuin pilottiaineistossa. Pohjois-Karjalan alueelta sen sijaan myös U-AlbKre-kokeiden tulokset löytyivät yhtä kattavasti kuin verikokeidenkin tulokset.

U-AlbKre-koee on laskennallinen suure, joka lasketaan U-Alb ja U-Krea kokeiden suhteena. Lisäksi tulosten ilmoittamisessa käytetään yleisesti \*, yli-, alle- < ja >-merkkejä osoittamaan mittaustarkkuuden alittavia tai ylittäviä lukuarvoja. Verikokeiden tulokset sen sijaan ilmoitetaan aina yksiselitteisesti lukuarvoina.

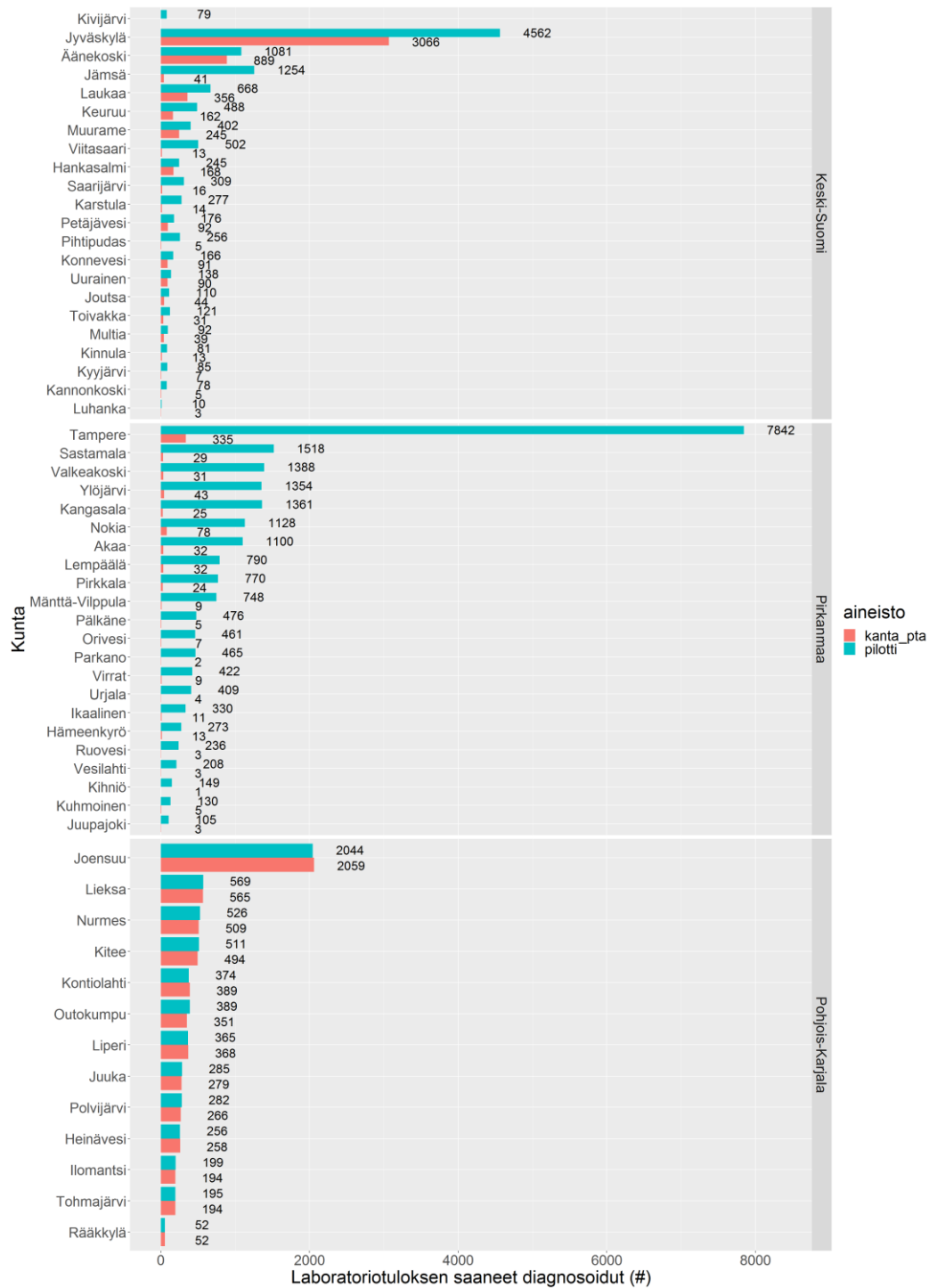


Kuva 3.26: Niiden potilaiden osuus PTA-aineiston potilaista, joille löytynyt P-Krea-tulos (%)



Kuva 3.27: Niiden potilaiden osuus PTA-aineiston potilaista, joille löytynyt U-albKre-kokeen tulos (%)





Kuva 3.28: Keski-Suomen, Pirkanmaan ja Pohjois-Karjalan hyvinvointialueilla kotikunnittain niiden henkilöiden lukumäärä, joille löytyy U-AlbKre-laboratoriotulos.

## Heikon laboratorioekokattavuuden syitä

Keskeisin syy todettuun PTA-aineiston huonoon laboratorioekokattavuuteen Pirkanmaan ja Keski-Suomen alueella on, että siellä käytetty potilastietojärjestelmä Pegasos käyttää laboratorioanalyysien tunnisteina ns. sisäistä avainkoodia Kuntaliiton virallisten koodien sijasta. Piilotaineiston tietopyyntö nojautui yksinomaan Kuntaliiton koodeihin, minkä vuoksi kaikki nämä analyysitulokset puuttuvat tutkitusta PTA-aineistosta.

Uranus- ja Esko-potilastietojärjestelmät eivät sisällä lainkaan laboratoriomodulia, vaan Uranuksen ja Eskon käyttäjät käyttävät erillistä MyLabin tai Fimlabin laboratorio-ohjelmaa. Mylabin laboratoriokokeet tallentaa PTA:oon Multilab-ohjelma. Fimlabin laboratoriokokeet tallentaa PTA:oon LKANTA-moduuli.

Seuraava taulukko kuvaa tietojärjestelmittäin, kuinka monta HbA1c-analyysiä vuonna 2019 PTA:oon tallennettiin, ja kuinka monessa analyysi identifioitiin Kuntaliiton koodilla 6128. Taulukko paljastaa systemaattisen virheen Pegasos-järjestelmässä. Tieto välitettiin CGI:lle, ja virhettä on alettu korjata. Korjattu Pegasos-versio on asennettavissa arviolta tammikuussa 2022. Näin ollen Pegasos-järjestelmästä laboratoriokokeita aletaan saada PTA:oon Kuntaliiton laboratorionimikkeistön koodeilla keväällä 2022.

Järjestelmä ja rekisterinpitäjä	HbA1c lkm	Koodi 6128 lkm	KL %
Abilita Hälsovård	7605	7605	100 %
Acute	6233	6233	100 %
DynamicHealth eArkisto	1130	1130	100 %
Effica eArkisto	452772	452251	100 %
Epic	12733	12733	100 %
LKANTA	20128	20128	100 %
Mediatri	60411	60411	100 %
Multilab	116364	116364	100 %
Pegasos	330716	0	0 %
SoftMedic	954	954	100 %
<b>Kaikki yhteensä</b>	<b>1009046</b>	<b>677809</b>	<b>67 %</b>

Pegasos-järjestelmää käyttivät seuraavat julkisen terveydenhuollon rekisterinpitäjät vuonna 2019 (taulukossa HbA1c-analyysien lkm). Kyseessä on samat rekisterinpitäjät, joiden alueilla todettiin myös matalia laboratoriokattavuuksia. Vastaava ongelma todettiin myös muissa laboratorioskoeanalyysissä (LDL, Krea,U-AlbKre).

Pegasos	330 716
Helsingin sosiaali- ja terveystoimi rekisterinpitäjä	82 959
Turun kaupungin sosiaali- ja terveystoimi rekisterinpitäjä	40 525
Tampereen sosiaali- ja terveystoimi rekisterinpitäjä	39 892
Kuopion kaupunki, rekisterinpitäjä	25 408
Vaasan kaupungin sosiaali- ja terveystoimi rekisterinpitäjä	9 622
Hyvinkään sosiaali- ja terveystoimi rekisterinpitäjä	8 657
Sastamalan seudun sosiaali- ja terveystoimi rekisterinpitäjä	7 242
Pohjois-Savon sairaanhoitopiirin kuntayhtymä rekisterinpitäjä	7 232
Keski-Satakunnan terveydenhuollon kuntayhtymä rekisterinpitäjä	6 693
Ylä-Savon SOTE kuntayhtymä, rekisterinpitäjä	6 616
Lohjan terveyskeskus rekisterinpitäjä	6 482

Perusturvakuntayhtymä Karviainen rekisterinpitäjä	5 678
Kirkkonummen perusturva rekisterinpitäjä	5 417
Sote kuntayhtymä Saarikka rekisterinpitäjä	5 037
Sosiaali- ja terveysvirasto, Naantalin kaupunki rekisterinpitäjä	4 528
Ylöjärven perusturvaosasto rekisterinpitäjä	4 415
Akaan perusturvatoimi rekisterinpitäjä	4 032
Keski-Uudenmaan sote-kuntayhtymä, rekisterinpitäjä	3 992
Raaseporin perusturva rekisterinpitäjä	3 975
Loimaan sosiaali- ja terveyspalvelukeskus rekisterinpitäjä	3 971
Paraisten sosiaali- ja terveystoimi rekisterinpitäjä	3 717
Loviisan sosiaali- ja terveystoimi rekisterinpitäjä	3 638
Kemijärven kaupunki rekisterinpitäjä	3 415
Pieksämäen kaupungin perusturva rekisterinpitäjä	3 341
Valkeakosken sosiaali- ja terveyskeskus rekisterinpitäjä	2 875
Wiitaunionin perusturva rekisterinpitäjä	2 802
Lempäälän terveyskeskus rekisterinpitäjä	2 599
Tornion terveyskeskus rekisterinpitäjä	2 394
Sisä-Savon terveydenhuollon kuntayhtymä rekisterinpitäjä	2 099
Huittisten perusturvakeskus, Huittisten kaupunki rekisterinpitäjä	2 071
Säkylän sosiaali- ja terveystoimi rekisterinpitäjä	1 852
Kalajoen perusturvakeskus rekisterinpitäjä	1 807
Kemin terveyskeskus rekisterinpitäjä	1 673
Järvenpään terveyspalvelut rekisterinpitäjä	1 525
Keminmaan terveyskeskus rekisterinpitäjä	1 405
Mustijoen perusturva rekisterinpitäjä	1 372
Posion kunta hyvinvointilautakunta rekisterinpitäjä	1 296
Sodankylän kunta rekisterinpitäjä	1 267
Inarin kunnan sosiaali- ja terveyslautakunta	1 255

Ylitornion terveyskeskus rekisterinpitäjä	993
Pelkosenniemen-Savukosken kansanterveystyön kuntayhtymä rekisterinpitäjä	821
Inkoon terveyspalvelut rekisterinpitäjä	802
Oulunkaaren kuntayhtymä rekisterinpitäjä	653
Kolarin kunta terveyskeskus rekisterinpitäjä	563
Siuntion terveyspalvelut rekisterinpitäjä	479
Tervolan terveyskeskus rekisterinpitäjä	465
Siilinjärven Sosiaali- ja terveyspalvelut rekisterinpitäjä	432
Kinnulan perusturvapalvelut rekisterinpitäjä	396
Pirkanmaan sairaanhoitopiirin kuntayhtymä rekisterinpitäjä	255
Eurajoen kunta, sosiaali- ja terveystoimi rekisterinpitäjä	69
Päijät-Hämeen hyvinvointikuntayhtymä rekisterinpitäjä	12
Länsi-Pohjan sairaanhoitopiiri rekisterinpitäjä	0

Laboratoriokokeiden puuttumiseen PTA:sta arvioitiin olevan muitakin syitä kuin Pegasoksesta löytynyt systeemivirhe.

Muiden syiden selvittämiseksi käytimme PTA:sta saatavaa "aggregoidut laboratoriokokeet"-aineistoa, josta löytyy kaikkien PTA:oon tallentuneiden laboratoriokokeiden nimet, numerosarjat ja laboratoriokokeen pyytäneen rekisterinpitäjän nimi. Teimme apuaineistossa laboratoriokokeen nimeen perustuvan tekstihaun ja kartoitimme, millaisilla numerosarjoilla ja nimillä diabetesrekisterissä tarvittavia laboratoriokokeita löytyi PTA:sta. Saimme selville, että PTA-aineiston laboratoriokattavuutta voivat pienentää myös muut satunnaiset virheet, joita tulee jatkossakin kartoittaa ja karsia yhteistyössä paikallisten toimijoiden kanssa. Osa virheistä voidaan oikaista laajentamalla jatkossa tietohakua aggregoidut laboratoriokokeet-aineistosta löytyneisiin laboratorionumeroihin. Osa virheistä on vielä mahdollisesti tunnistamatta niin tarkasti, että korjausehdotusta voitaisiin esittää.

Jotkut yleisesti käytössä olevat, ja tietopyynnöstä puuttuneet koodit saattavat selittää osan alueellisista eroista kattavuudessa. Tämä voidaan jatkossa korjata lisäämällä tietopyyntöön toistaiseksi laboratoriokoodeja. Jatkossa tavoite tulee kuitenkin olla, että kansallisesti käytetään Kuntaliiton nimikkeistön kansallisia koodeja kattavasti, ja PTA:sta laboratoriokokeet ovat niiden avulla löydettävissä yksiselitteisesti.

- P-LDL-kolesteroliaineistoon tulee jatkossa hakea plasmasta paastossa ja ei-paastossa otetut LDL- kokeet: fP-LDL (4599) ja P-LDL (6433), kun nyt haettiin vain paasto LDL 4599. Tämä oli tietopyynnössä tehty määrittelyvirhe, joka oikaistaan.
- HbA1c arvot löytyvät kattavasti käyttämällä kuntakoodia B-HbA1c (6128), mutta sen lisäksi yleistyy vieritestien käyttö, jotka kannattaa jatkossa lisätä tietopyyntöön c-HbA1c (tai vieriHbA1c, 6591). Jonkin verran on vielä käytössä HbA1c,% (1560), vaikka kansallinen päätös prosenttimuodossa ilmoitetusta HbA1c:sta tehtiin jo 2010.

- eGR-laskemista varten tietopyyntöön haetaan P-krea (4600) lisäksi paastokoe fP-krea (2142) sekä s-krea (2143).
- U-albkre tietopyyntöön haetaan virallisen 4511 koodin lisäksi jostain syystä hyvin yleisesti käytetty epävirallinen koodi 9522.

U-AlbKre-arvoissa esiintyy paljon <-, >- tai \*-merkkejä ja alle-sanoja, mikä saattoi selittää poikkeavan suurta tietojen puutteellisuutta. Jatkossa arvoksi poimitaan merkkiä seuraava lukuarvo eli mittaustarkkuuden alarajan lukema.

Havaitsimme myös virheellisesti käytettyjä koodeja, esim 4122 on selkäydinnäytteen koriongonadotropiinin laboratorionimikkeistön koodi, mutta sen selityksenä käytettiin P-LDL-kolesterolia. Näistä virheistä ilmoitetaan rekisterinpitäjille, mutta näitä kuntakoodeja ei oteta mukaan diabetesrekisteriin jatkossa.

Alhaista kattavuutta saattavat selittää myös paikallisesti käytetyt labrakoodit, joiden taustalla voivat olla toimipaikkojen käyttötarpeet, jotka eivät liity hoidolliseen käyttöön. Paikallisia koodeja käytetään esimerkiksi toiminnanohjaukseen, kuten laboratoriodien työjonoihin, sisäiseen ja ulkoiseen laskutukseen, alihankkijaketjujen sanomaliikenteeseen ja toiminnanohjausraportointiin. Näiden osuutta pystytään arvioimaan vasta, kun suuren kattavuusvauksen aiheuttaneet virheet on korjattu.

## Esimerkkejä kansallisesti skaalatusta diabetesrekisteristä saatavista hoitotasapainoa kuvaavista laatuindikaattoreista

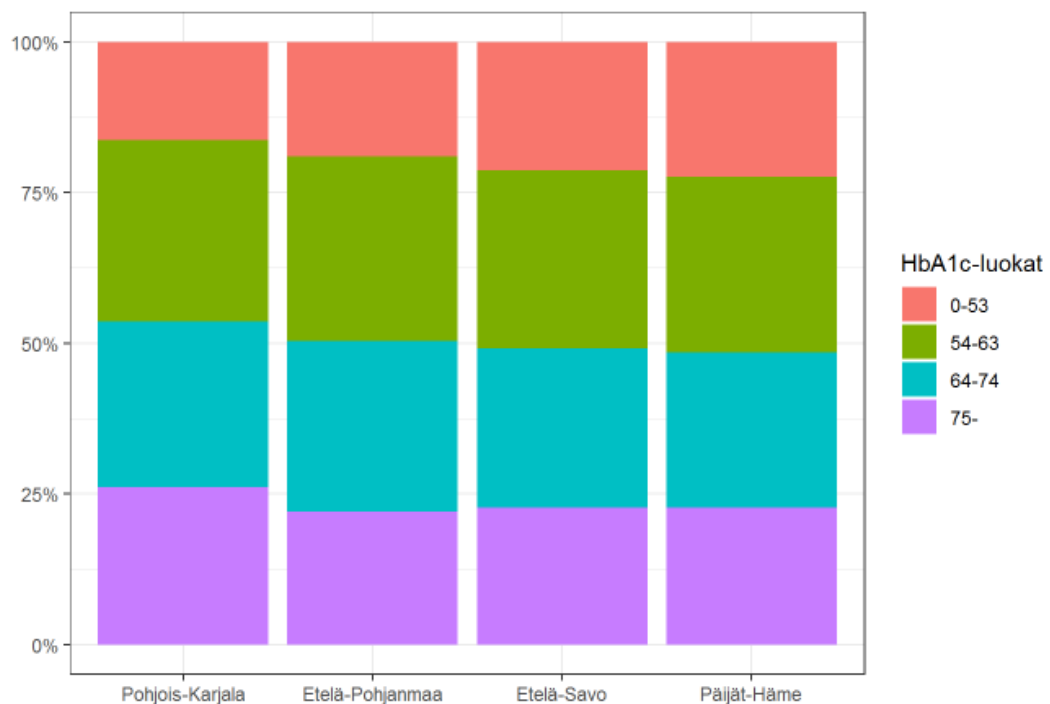
Alla näytetään esimerkkejä niistä laatuindikaattoriraporteista, joita on jatkossa tarkoitus tuottaa hyvinvointialueille laboratoriarvoihin perustuen. Raporttien tekninen toteutus ja visualisointi on oleellisen tärkeä kehittämiskohde pysyvää raportointia suunniteltaessa. Tärkeää on saada raporteille näkyviin myös tulosten muutos suhteessa aikaan ja suhteessa vertailualueiden tulosten muutokseen sekä yleiseen trendiin. Esimerkkikuvissa esitetään hoitotasapaino vain niillä hyvinvointialueilla, joissa diagnoosien saaneiden kattavuus oli yli 80 % summapopulaation potilaista ja joissa laboratorioskokeiden kattavuus on yli 80 % diagnoosin saaneista diabeetikoista. Lähde on PTA-datasta muodostettu analyysiaineisto. Mukaan on otettu vain henkilön viimeisin mittaustulos viimeisen kahden vuoden aikana (2019–2020) elossa olevilta. Tutkimukset, joista tutkimustulosarvoa ei löydy, on poistettu aineistosta. Aluetiedot perustuvat henkilön kotikuntaan. Mukana ovat vain ICD-10-diagnoosin saaneet henkilöt ja tutkimusnimikkeellä 6128 (HbA1c) ja 4600 (P-Krea/eGFR) poimitut laboratoriovastaukset.

Kuvasta 2.29 nähdään, että 25-74-vuotiaista tyyppin 1 diabeetikoista noin puolet on huonossa tai erittäin huonossa sokeritasapainossa (HbA1c >64mmol/mol) ja lähes 25 % erittäin huonossa hoitotasapainossa (HbA1c yli 75mmol/mol). Rekisteridata tuo esiin selkeän hoitovajeen ja kehittämiskohteen. Hyvinvointialueiden välillä on jonkin verran eroja hoidon laadussa. Kuvan 3.20 perusteella hyvinvointialueella voidaan arvioida, kuinka monen potilaan hoitoon tulisi puuttua, jotta erittäin huono tasapaino saataisiin korjaantumaan. Näitä potilaita on noin 700–800 kullakin hyvinvointialueella. Tämä tieto auttaa resurssien kohdentamisessa, kehittämiskohteen valinnassa ja haasteen suuruuden arvioinnissa.

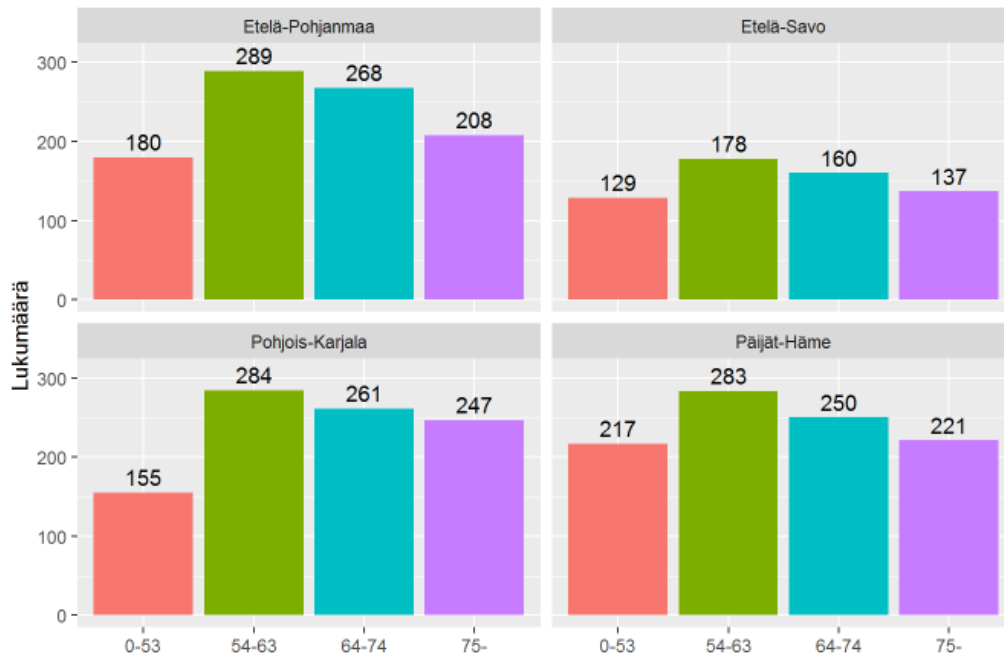
Kuvasta 3.21 nähdään, että tyyppin 2 diabeetikoista on suhteellisesti paljon pienempi osuus huonossa tai erittäin huonossa hoitotasapainossa kuin tyyppin 1 diabeetikoista. Kuitenkin kuva 3.32 paljastaa, että erittäin huonossa tasapainossa olevia tyyppin 2 diabeetikoita on lukumääräisesti enemmän kuin erittäin huonossa tasapainossa olevia tyyppin 1 diabeetikoita. Tyyppin 2 diabeetikoiden suuren lukumäärän vuoksi raportit antavat selkeän viestin siitä, että resursseja tulisi siirtää rutiininomaisesta kaikkien potilaiden seulonnasta erittäin huonossa tasapainossa olevien potilaiden hoitoon.

Kuvasta 3.33-3.36 nähdään, että diabeteksen munuaistaudin (eGFR<60) prevalenssi on noin 10 % kummassakin diabetestyyppissä. Nämä potilaat ovat erittäin korkean riskin potilaita eli kustannusvaikuttavinta olisi suunnata ennalta ehkäisevä glukoositasapainon, verenpainelien ja LDL-kolesterolin hoito näihin potilaisiin, jotta mahdollisimman moni välttäisi kalliin loppuvaiheen munuaistaudin tai sydänsairauden. Näitä potilaita on alle 1000 kullakin hyvinvointialueella.

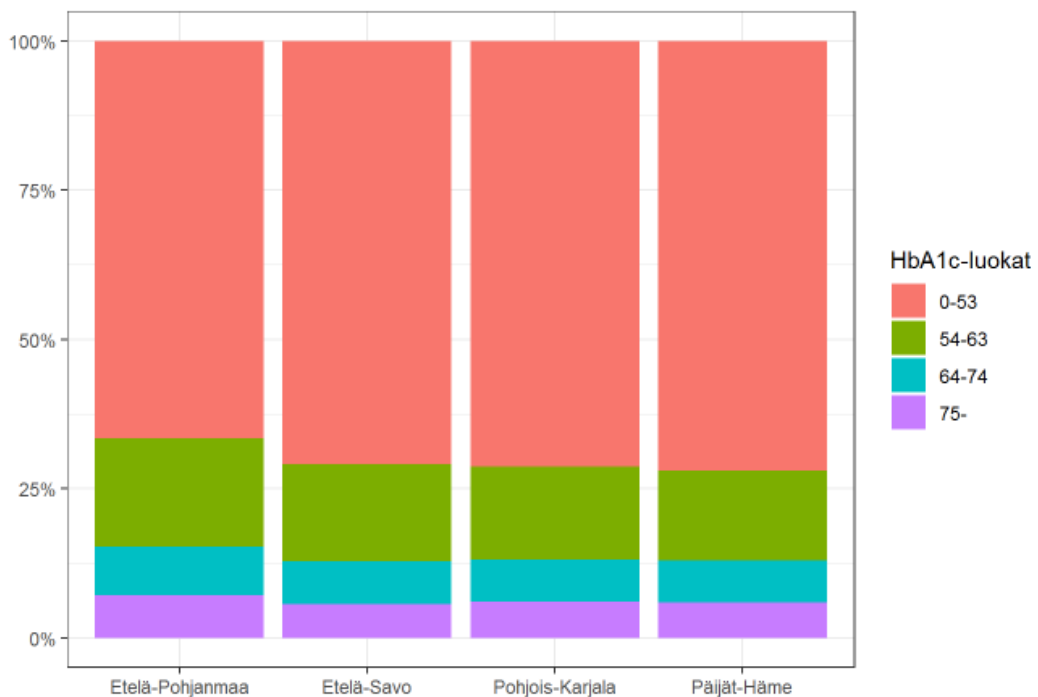
Kuvasta 3.37–3.38 nähdään, että LDL-kolesteroli on liian korkea (yli 2,6 mmol/l) 25–30 % sekä tyyppin 1 että tyyppin 2 diabeetikoista. Huonossa LDL-kolesterolitasapainossa olevien potilaiden lukumäärä on 2500–3000 kullakin hyvinvointialueella. Erittäin korkean riskin potilailla LDL-kolesterolitavoite on vielä tiukempi. LDL-kolesterolitasossa havaitaan selvä hoitovaje etenkin tyyppin 2 diabeetikoiden kohdalla, joilla on lisääntynyt riski saada sepelvaltimotauti. Huonossa hoitotasapainossa olevien potilaiden määrä on huomattavasti suurempi kuin esimerkiksi huonossa sokeritasapainossa olevien potilaiden määrä.



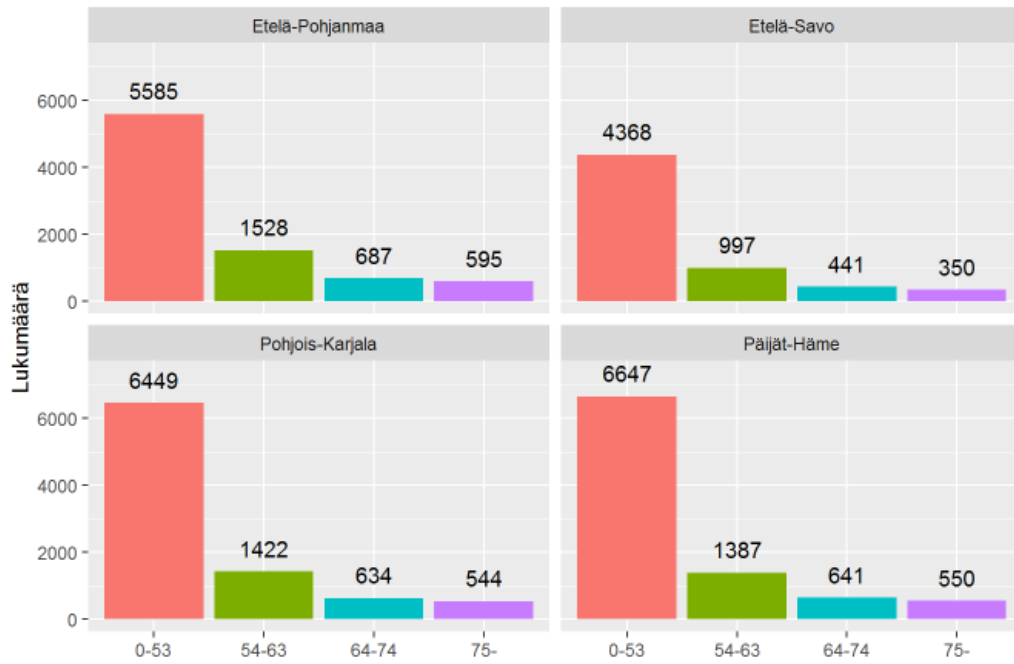
Kuva 3.29: HbA1c-arvot sokeritasapainoluokittain ja hyvinvointialueittain 25-74-vuotiailla tyyppin 1 diabeetikoilla.



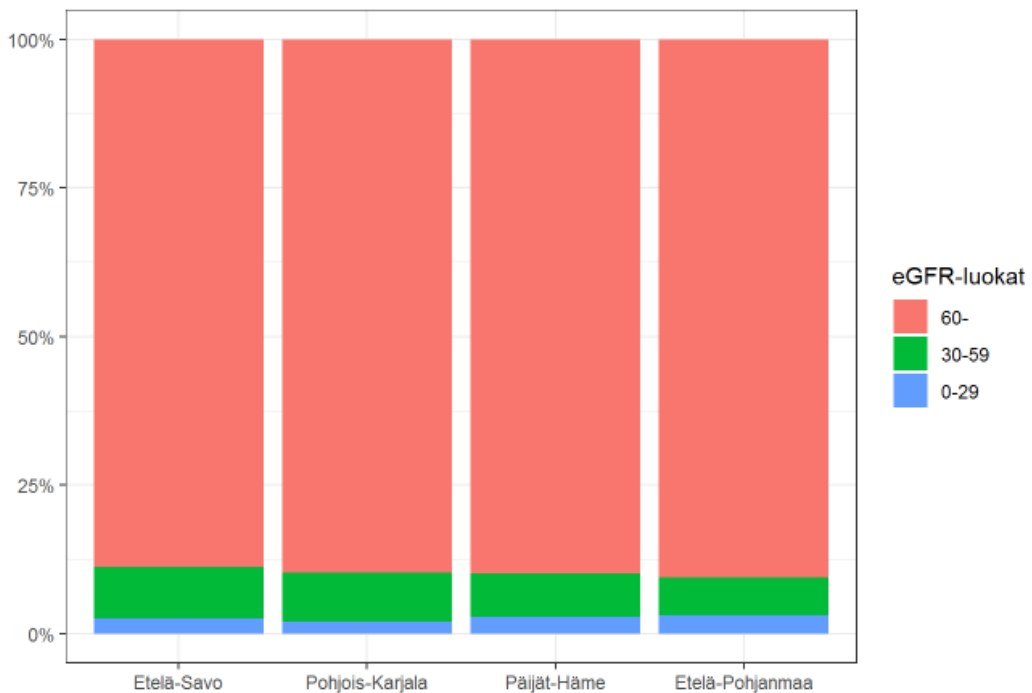
Kuva 3.30: HbA1c-kokeessa käyneiden 25-74-vuotiaiden tyypin 1 diabeetikoiden lukumäärät sokeritasapainoluokittain ja hyvinvointialueittain.



Kuva 3.31: HbA1c-arvot sokeritasapainoluokittain ja hyvinvointialueittain 25-74-vuotiailla tyypin 2 tai muu diabeetikoilla.

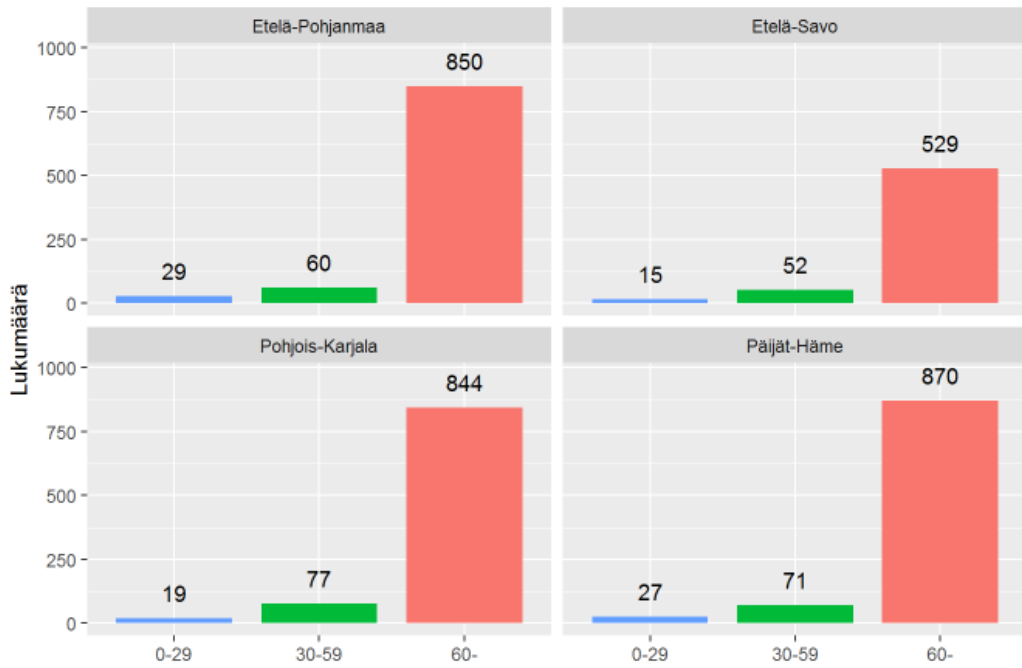


Kuva 3.32: HbA1c-kokeessa käyneiden 25-74-vuotiaiden tyypin 2 tai muu diabeetikoiden lukumäärät mittausluokittain ja hyvinvointialueittain.

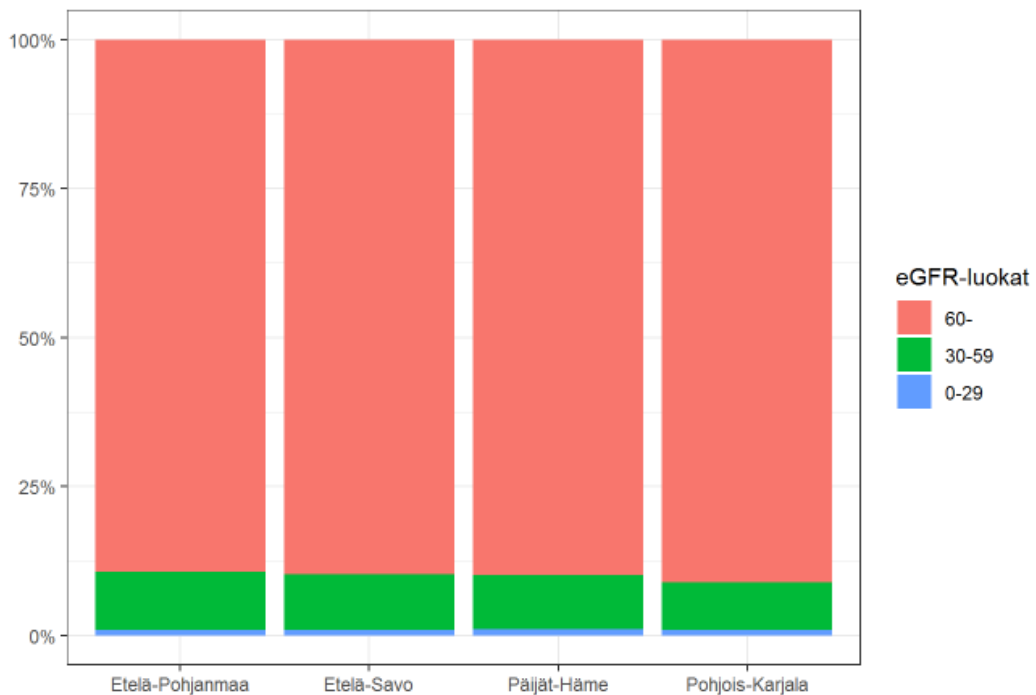


Kuva 3.33: eGFR-arvot (P-Krea-arvoista laskettu) luokittain ja hyvinvointialueittain 25-74-vuotiailla tyypin 1 diabeetikoilla.

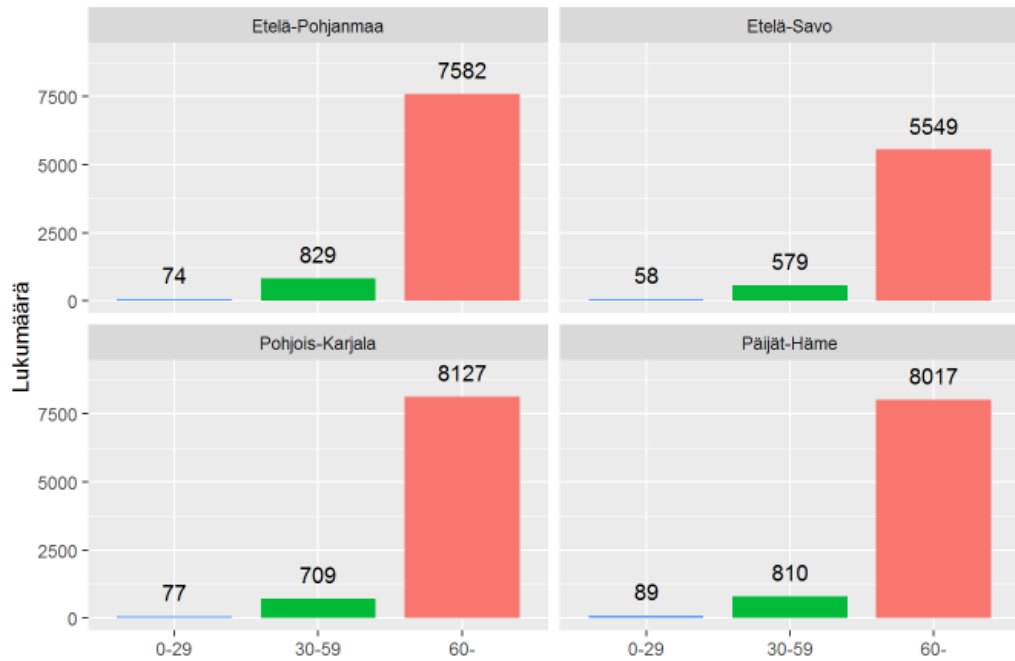




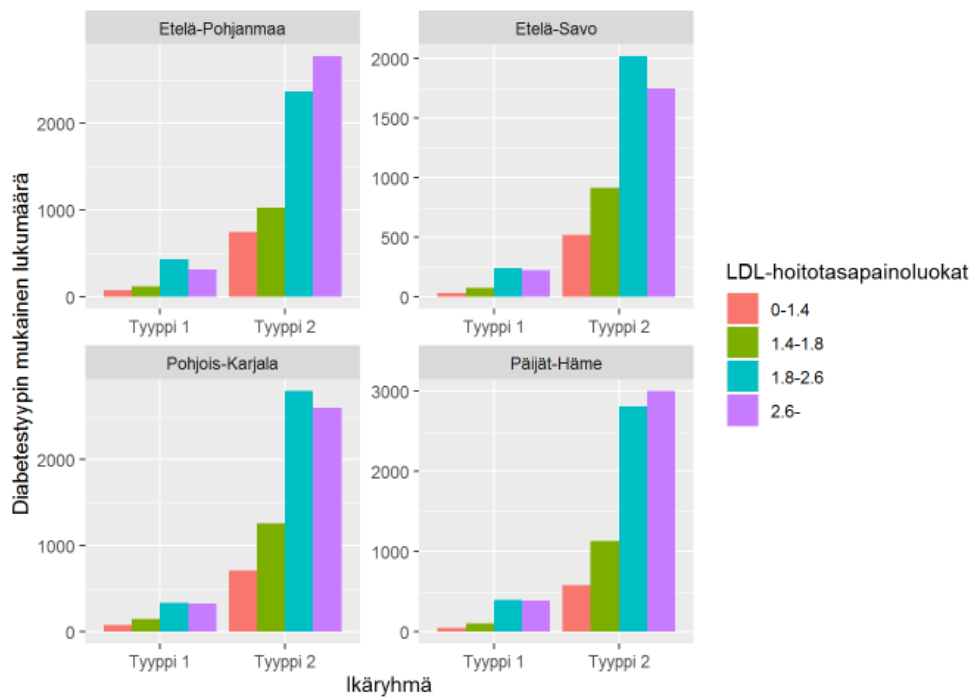
Kuva 3.34: P-Krea-kokeessa käyneiden 25-74-vuotiaiden tyypin 1 diabeetikoiden lukumäärät eGFR-luokittain ja hyvinvointialueittain.



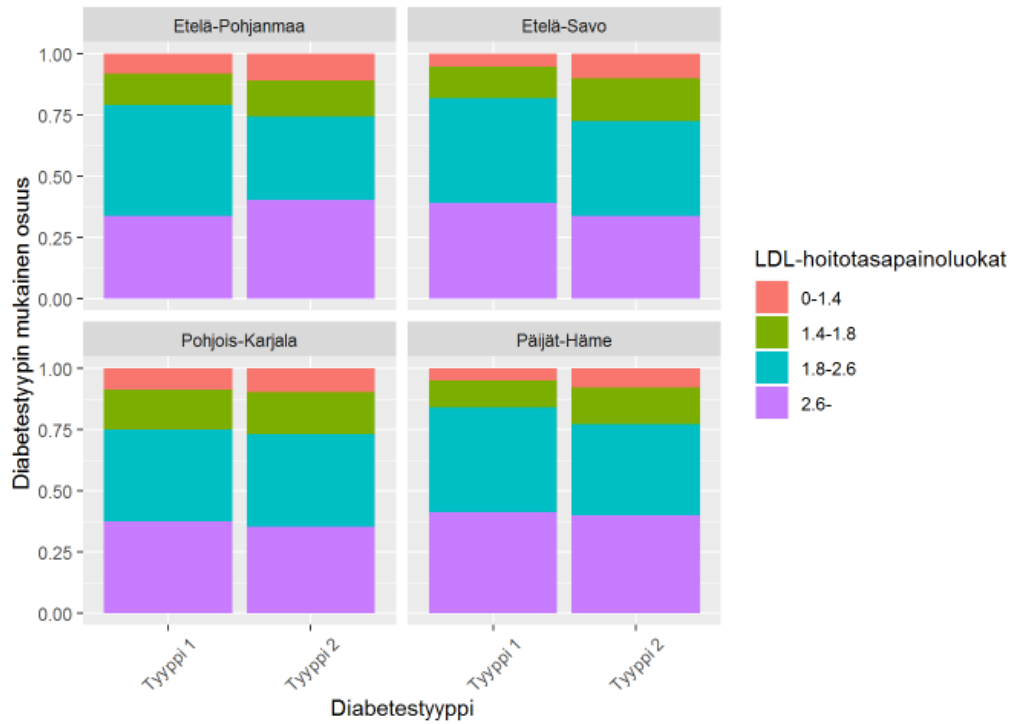
Kuva 3.35: eGFR-arvot (P-Krea-arvoista laskettu) luokittain ja hyvinvointialueittain 25-74-vuotiailla tyypin 2 tai muu diabeetikoilla.



Kuva 3.36: P-Krea-kokeessa käyneiden 25–74-vuotiaiden tyypin 2 tai muu diabeetikoiden lukumäärät eGFR-luokittain ja hyvinvointialueittain.



Kuva 3.37: Potilaiden lukumäärät LDL-kolesterolin hoitotasapainoluokittain ja diabetestyypeittäin neljällä tutkitulla hyvinvointialueella.



Kuva 3.38: Potilaiden suhteelliset osuudet LDL-kolesterolin hoitotasapainoluokittain ja diabetestyypeittäin neljällä tutkitulla hyvinvointialueella.

# 4 Havainnot ja johtopäätökset

## Potilaiden tunnistaminen

Diabetespotilaiden tunnistamiseksi käytettävissä oli kolme erillistä tietolähdettä:

1. Potilastiedon arkiston (PTA) ICD-10-diagnoositiedot (syy ja ulkoinen syy)
2. Reseptitietokannan (eResepti) lääkemääräykset ja -ostot: diabeteslääkkeet tunnistetaan ATC-luokasta A10
3. Lääkkeiden erityiskorvausoikeusrekisteri (etuusrekisteri): erityiskorvausoikeudet YE103 ja AE215.

Lisäksi validoinnin tukena oli käytettävissä THL:n Hoitoilmoitusjärjestelmästä

1. erikoissairaanhoidon (ESH) Hilmo-tietokantaan kirjatut ICD-10-diagnoosit vuosilta 1998–2020, ja
2. perusterveydenhuollon (PTH) AvoHilmo-tietokantaan vuoteen 2020 mennessä kertyneet ICD-10- ja ICPC 2-diagnoosit. ICPC 2-diagnoseja emme potilaiden lopullisessa tunnistamisessa kuitenkaan käyttäneet, koska nämä usein edustavat vasta diabetesepäilyä.

PTA:n ICD-10-diagnoosin perusteella todennettiin 402000 vuoden 2021 alussa elossa olevaa diabetespotilasta. Tämä on hieman vähemmän kuin mitä A10-reseptipotilaita on (445000). Toisaalta luku on merkittävästi suurempi kuin Hilmojen (esh + pth) ICD-10-diagnoosien perusteella päätelty potilasjoukko, 293000. Hilmon pienempään lukuun voi vaikuttaa ajallinen raja: TerveysHilmo sisälsi lähinnä käyntejä vuosilta 2018–2020. Vaikka AvoHilmo on käynnistynytkin jo vuonna 2010, on valtaosa kirjatusta käynneistä vuosilta 2019–2020 (ks. [luku 2 Aineisto ja menetelmät](#)).

Sosiaali- ja terveydenhuollon rekisteripitäjät ja niiden käyttämät asiakas- ja potilastietojärjestelmät (APTJ) ovat liittyneet Kanta-palveluihin käytännössä vuodesta 2014 lähtien. Peräti 170 julkisen terveydenhuollon 179:stä rekisterinpitäjästä oli liittynyt Kantaan vuoden 2016 alkuun tultaessa. Tämän jälkeen liittyneistä rekisterinpitäjistä vain yksi suuri rekisterinpitäjä, Pohjois-Karjalan sosiaali- ja terveyspalvelujen kuntayhtymä, liittyi vuoden 2017 alussa. Voitaneen arvioida, että vuoteen 2018 tultaessa useimmat julkisen terveydenhuollon rekisterinpitäjät ovat jo liittyneet Kantaan, ja diagnoosien ICD-10-koodit löytyvät siten PTA:sta.

Ahvenanmaata ei ole veloitettu tallentamaan potilastietoja Kantaan, vaikka sähköinen resepti on käytössä. Ahvenanmaan PTA-kirjaukset ovatkin hyvin vajavaiset. Tämän vuoksi keskeytimme analyysissä pääasiassa Manner-Suomeen.

Potilastieto- ja apteekkijärjestelmät liittyivät sähköiseen eResepti-järjestelmään 2010-luvun aikana siten, että jo vuonna 2014 reseptikanta on varsin edustava (ks. luku 2). Vuodesta 2017 lähtien liittyminen on ollut pakollista. Potilaiden tunnistamisessa käytimme mahdollisimman herkkää sääntöä: henkilö identifioitiin lääketietojen perusteella diabetespotilaaksi, jos hänelle oli kirjoitettu yhdenkin diabeteslääkkeen resepti. Lääkinnällisiä alaryhmiä muodostettaessa käytimme kuitenkin kriteerinä lääkemääräyksen sijasta potilaan viimeisimmän vuoden aikaisia diabeteslääkkeiden ostoja.

Etuusjärjestelmästä poimitut erityiskorvattavuusoikeudet identifioivat 363000 potilasta, joille oli vuosina 2010–2020 myönnetty joko insuliinilääkkeen korvaus (koodi YE103) tai ei-insuliinikorvaus (AE215), tai kummatkin. Tietopyynnön muotoilun vuoksi aineisto ei sisältänyt ennen vuotta 2010 tehtyjä YE103-etuuspäätöksiä. Nähtävästi kuitenkin kaikille vuonna 2016 elossa olleille, aiemmin jo myönteisen YE103-päätöksen saaneille on myönnetty vuonna 2016 myös AE215 etuus, joten aineisto on tältä osin kattava.

PTA:oon kirjattujen diagnoosien perusteella muodostetusta potilasjoukosta puuttuvat selvästi ainakin Ahvenanmaan potilaat. Myös Pohjanmaalla, Pohjois-Pohjanmaalla ja Kainuussa on PTA:n mukaan melko alhaiset prevalenssit, 5–6 %. Reseptikannan perusteella laskien prevalenssi vaihtelee 7–10 %. Uudenmaan PTA-diagnoosin perusteella arvioitu prevalenssi on myös alhainen, 6 %, mutta se voi osin selittyä myös nuoremmalla ikärakenteella.

Diagnoosin toteamispäivä on ongelmallinen. Mitä ilmeisimmin se on useimmiten diagnoosin kirjaamispäivä eikä sairauden ensimmäinen toteamispäivä. Ensimmäiset lääkeresepit on kirjattu reseptitietokantaan kohtuullisen kattavasti vasta vuodesta 2013 tai 2014. Erityiskorvatavuuskoodi 215 otettiin käyttöön vasta vuonna 2016. Koodi 103 on otettu aiemmin käyttöön, mutta tässä aineistossa ensimmäinen vuosi on 2010 huonosti muotoillun tietopyynnön vuoksi. Kuolleisuusanalyysissä käytimme sairauden havaitsemisen alkuna päivämäärää, jolta ensimmäinen merkintä potilaan diabeteksestä oli olemassa minkä tahansa tietolähteen mukaan.

Kolmen erillisen tietolähteen käyttö parantaa identifioinnin herkkyyttä eli sensitiivisyyttä, kun valitaan summapopulaatio, eli ne potilaat, jotka esiintyvät yhdessäkin näistä tietolähteistä. Tämä oli meidän alustavien analyysiemme lähtökohta. Toisaalta tietolähteitä voidaan käyttää identifioinnin tarkkuuden eli spesifisyyden (diagnoosin varmuusasteen) määrittämiseen rajautumalla analyysissä esimerkiksi vain niihin potilaisiin, jotka kuuluvat kaikkiin tietolähteisiin. Samaa sensitiivisyys-spesifisyys-akselia voidaan jakaa analyysissä vielä hienojakoisemmin rajautumalla esimerkiksi lääketiedoissa lääkemääräysten sijasta ostoihin tai potilaan viimeaikaiseen lääkeadherenssiin, diagnooseissa näiden potilaskohtaiseen esiintymisfrekvenssiin tai -ajankohtaan, jne.

Kolmen tietolähteen summapopulaatio tuntuu poimivan hyvin kaikki potentiaaliset potilaat ainakin, jos vertaamme käyttämiimme Hilmo-referensseihin. Vain 0,7 % summapopulaation ja Hilmon muodostamasta yhdisteestä oli sellaisia, jotka löytyivät vain jommasta kummasta Hilmosta. Hilmot yhdessä kattoivat noin 60 % summapopulaatiosta. Tarkempaa analyysiä tarvitaan vielä selvittämään summapopulaation diabetessairastavuuden varmuusastetta.

Diabeteksen luokittelu tyyppin 1 ja 2 sairauksiin edellyttää ICD-10-diagnoosia. Valitettavasti alueelliset analyysimme osoittavat diagnoosien alikirjaamista joillain alueilla. Kun vertasimme alueittain PTA:sta löytyvien ICD-10-diagnoosien esiintyvyyttä reseptitietokannan diabeteslääkemääräyksiin, havaitsimme vajetta monilla hyvinvointialueilla (Kainuu, Pohjois-Pohjanmaa, Länsi-uusimaa, Lappi, jne; ks. kuva 3.4a [luvussa 3 Tulokset](#)). Emme löytäneet systeemistä syytä ilmiöön. Asiaa on kuitenkin vielä tutkittava tarkemmin – tai viime kädessä tehtävä suosituksia hoitojen kirjauskäytänteisiin.

Toisaalta lääkevertailu voi paljastaa myös ylidiagnosointia, jolloin alueen aktiivinen suhtautuminen diabeteksen hoitoon näkyy diabetespäilyjen kirjaamisena ennen lääkityksen määrittämistä (kuva 3.4b). Diabetesdiagnoosin esiintyminen ilman diabeteslääkemääräystä voi lisäksi selittyä muun muassa laitoshoidolla tai diabeteksen paranemisella elämäntapamuutosten myötä.

### Diagnoosit

Diagnoosien ICD-10-koodit ovat keskeisiä paitsi potilaiden tunnistamisen niin myös oheissairastavuuden ja diabeteksen aiheuttamien komplikaatioiden kannalta.

Yksi potilastiedon arkiston eduista on, että *a priori* se kattaa myös yksityistä terveydenhuoltoa Hilmoa paremmin. Sairastavuuden määrittelyssä tavoite on tukeutua pelkästään PTA:n diagnooseihin. Tässä pilotissa meillä oli käytettävissä myös THL:n TerveysHilmo validointia varten.

Tietopyynnössä esiintyneen virheen vuoksi alkoholin ja huumeiden käyttöön liittyvät PTA-diagnoosit puuttuivat aineistosta kokonaan. Toisaalta taas ien- ja juuri-infektiot (K04, K05) eivät löydy TerveysHilmosta, minkä vuoksi näiden osalta vertailua ei ollut tehtävissä. Emme ehtineet tehdä vastaavaa vertailua käyttäen AvoHilmoa.

Vertasimme PTA:oon kirjautuneita ICD-10-diagnooseja TerveysHilmon diagnooseihin sekä aggregaattitasolla (lukumäärinä) että parivertailuna. Parivertailu tarkoittaa sitä, että potilaalle Hilmon hoitojaksolla kirjattuja yksittäisiä ICD-10-diagnooseja haetaan PTA:n ICD-10-diagnosien joukosta sellaisella kirjausajalla, joka osuu Hilmo-hoitojaksolle.

Useiden keskeisten diagnoosien kohdalla aggregaattitason vertailu TerveysHilmoon antaa tyydyttäviä tuloksia.

Ehdimme tehdä vain joitain pistokokeen omaisia parivertailuja tietyillä diagnooseilla (E10, I25, Z49.1, Z49.2, Z94.0). Tältä osin tulokset eivät olleet tyydyttäviä: pystyimme tunnistamaan PTA:sta vain noin puolet Hilmon hoitojaksoilla kirjatusta diagnosoineista.

Karkeistimme parivertailua vielä siten, että tarkastelimme asiaa potilaan sairastavuustasolla: löytyvätkö esimerkiksi samat dialyysi- ja munuaissiirtopotilaat Hilmosta ja PTA:sta? Tuloksena oli, että 68 % (1646/2428) dialyysipotilaista oli kirjattu Z49-diagnosilla sekä Hilmoon että PTA:han. Vastaavasti 86 %:lle (1396/1632) oli kirjattu munuaissiirtodiagnoosi X94.0 sekä Hilmoon että PTA:han.

Huomattakoon, että dialyysihoito on erikoissairaanhoidossa laskutettavaa toimintaa, joten diagnoosi puuttunee hyvin harvoin kirjaamattomuuden vuoksi. Selvitettäväksi jää, voiko tässä olla kyse teknisestä virheestä jossakin prosessoinnin vaiheessa.

Diagnoosien kirjautuminen PTA:han edellyttää vielä paljon jatkotutkimusta. Emme löytäneet mitään systemaattista selitystä esimerkiksi edellä kuvatulle huonolle Hilmon ja PTA:n yhteensopivuudelle.

Havainnot viittaavat siihen, että joskin prevalenssiestimaatteja saanee aineistosta riittävän luotettavasti, sairauksien ilmaantuvuuslukujen (insidenssien) arviointia varten yhteensopimattomuudet pitää pystyä selvittämään ja mahdolliset systeemiset virheet korjaamaan. Hilmon lisäksi validoinnissa voitaisiin käyttää iskeemisten sydänsairauksien rekisteriä ja munuaistautirekisteriä.

Havaitsimme suurta vaihtelua diabetespotilaiden vakioidussa kuolleisuudessa eri hyvinvointialueiden kesken. Kuolleisuuden luotettava vertailu edellyttää kuitenkin huolellista diabeteksen vaikeusasteen ja oheissairastavuuden hallintaa (vakiointia) sekä mahdollisimman tarkkaa ja alueellisesti yhtenäisellä tavalla mitattua tietoa sairauden syntyhetkestä. Jatkossa on arvioitava aineistoon poimittavien oheissairauksien diagnosit myös kuolleisuuden vakioinnin kannalta: aineistosta olisi kyettävä laskemaan esimerkiksi lisäsairauksia eli komorbiditeettia kuvaavaa Charlson-indeksi, jota siten voitaisiin käyttää yksittäisenä oheissairastavuutta vakioivana muuttujana. Pilottiaineistoamme kattavammat ja ajallisesti tarkemmat diabeteksen kestoja mittaavat diagnositiedot ovat välttämättömät.

## Toimenpiteet

Niin diabeteksen aiheuttamien komplikaatioiden kuin hoidollisten alaryhmienkin muodostamisen kannalta tieto potilaalle tehdyistä tietyistä toimenpiteistä on tärkeää.

Toimenpiteiden kirjautumisen validointiin käytimme TerveysHilmon aineistoa.

Radiologiset toimenpiteet puuttuvat käytössämme olleesta PTA:n toimenpiteiden taulusta. Joko radiologian tietojärjestelmät eivät kirjaa näitä rakenteisesti, tai ne kirjataan jotenkin poikkeavalla tavalla, minkä vuoksi ne eivät poimiudu PTA Tietoallas -toimenpiteiden tauluun kuten muiden järjestelmien kirjaamat toimenpiteet. Asia vaatii tarkempaa selvittämistä.

Tietoallas-mallinnuksen mukaisesti toimenpiteellä voi olla useita tekijöitä. Toimenpiteen tekijät on kirjattu omaan tauluunsa, joka sisältää myös tiedon organisaatiosta, jossa toimenpide tehtiin, sekä toimenpiteen suorituspäivästä (tai organisaatioista ja päivistä).

Toimenpiteen ajankohta on luonnollisesti tärkeä tieto. Tämä on kuitenkin käytettävissä vasta vuodesta 2020 lähtien. Meille ei selvinnyt, miksi tieto puuttuu aiemmilta vuosilta.

### Laboratoriotulokset

Potilastiedon arkistosta poimittavia laboratoriotuloksia arvioitiin hyvinvointialueittain ilmoitettavina kattavuusarvoina eli kuinka monelta PTA-diabetesdiagnoosin E10–E14 saaneelta potilaalta löytyy diabetesrekisterin ydinindikaattoreihin kuuluva laboratoriotulokset viimeisen 2 vuoden ajalta. Mukana oli useita hyvinvointialueita, joilla keskeisten indikaattorien (HbA1c ja LDL-kolesteroli) kattavuudet olivat yli 80 %. Joukossa oli kuitenkin myös alueita, joissa kattavuus oli huomattavasti matalampi. Lisäksi arvioitiin laboratoriotuloksissa käyneiden potilaiden vastaavuutta ja löydettyjen laboratoriotulosten määrää pilottimaakunnissa. Pilottialueissa Siun soten potilastietojärjestelmästä löytyivät pääsääntöisesti samat potilaat ja laboratoriotulokset kuin PTA:sta. PTA:sta oli mahdollista poimia myös yksityissektorilla määritetyt kokeet. Molemmilla tavoilla arvioitiin, että PTA on luotettava ja käyttökelpoinen laboratoriotulosten tietolähde, kunhan joitakin korjauksia vielä tehdään.

Alueellinen jako perustui analyysissämme potilaan viimeisimpään kotikuntaan. Muuttoliikkeen yms. vuoksi alueelliset vertailut eivät ole näin ollen täysin tarkkoja, mutta arviomme mukaan riittävästi suuntaa antavia alueellisen kattavuuden ja tulosten vertaamiseen.

Diabetesrekisterin laboratoriotuloksiin perustuvia ydinindikaattoreita ovat HbA1c ja LDL-kolesteroli. Lisäksi diabeetikoiden sydän- ja verisuonisairauksien riskitason mukaisesti alaryhmäjakoihin tarvitaan eGFR- ja U-AlbKre-kokeita. Näiden laboratoriotulosten tuloksia tutkittaessa havaitsimme, että Pegasos-sairauskertomusjärjestelmää käyttäviltä rekisterinpitäjiltä ei löydy ydinindikaattorilaboratoriotuloksia Kuntaliiton virallisilla tutkimuskoodilla, koska Pegasos-järjestelmä kirjaa potilasasiakirjoihin virallisen koodin sijasta järjestelmän sisäisen analyysin avainnumeron. CGI on ilmoittanut muuttaneensa ohjelman toimintaa 9.12.2021 julkaisemassaan Pegasos-versiossa 9.2.18. Uuden ohjelmaversioon arvioidaan olevan kaikkien asiakkaiden käytössä viimeistään helmikuussa 2022. On kuitenkin edelleen mahdollista, että asiakkaat käyttävät joidenkin laboratorioanalyysien tunnisteen ns. paikallisia tutkimuskodeja Kuntaliiton virallisten koodien sijasta. Asian korjaamiseksi lopullisesti pitää myös nämä päivittää ohjelman rekistereihin, ja tämä täytyy tehdä paikallisesti. Ongelma tuskin kuitenkaan koskee vakiintuneita analyysiejä kuten HbA1c tai LDL-kolesteroli. Asiaa on kuitenkin seurattava aktiivisesti jo vuoden 2022 alkupuoliskolla. On syytä huomata, että takautuvasti tämä ohjelman muutos ei tietenkään korjaa PTA-aineistoa. Jos rekisteri halutaan takautuvasti päivittää, niin silloin on etsittävä laboratorioanalyysiejä tutkimuskoodin lisäksi myös tekstimuotoisesta nimikekentästä.

Diabetesrekisterin kannalta keskeiset laboratoriotulokset HbA1c ja LDL-kolesteroli ovat sellaisia, että Kuntaliiton tutkimuskoodit ovat käytännössä sataprosenttisesti käytössä luukuunnottamatta Pegasosta. Sen sijaan U-AlbKre- ja cU-Alb-analyysissä on käytössä myös paikallisia tutkimuskodeja (ks. jäljempää).

Käytettävissämme oli PTA:han tallennetuista, yli 721 miljoonasta laboratorioanalyysistä muodostettu aggregaattitaulu (ns. "aggregoidut\_labrat"), joka oli keskeinen selvittelyn työväline. Se sisälsi lukumäärätiedot kaikista julkisessa terveydenhuollossa tehdyistä analyysistä rekisterinpitäjän, tietojärjestelmän, tutkimusnimikkeen ja nimikekoodin suhteen luokiteltuina. Tämän taulun avulla pystyimme arvioimaan, millä eri tavoilla järjestelmät ja rekisterinpitäjät kirjaavat analyysituloksiaan PTA:han. Se mm. paljasti, kuinka yleistä on poiketa Kuntaliiton nimikkeistöstä. Vastaavanlainen aggregaatti on syytä tehdä vuoden 2022 aikana kirjattavista laboratorioanalyysistä jo vuoden alkupuoliskolla, jotta esimerkiksi voidaan varmistua edellä kuvatun Pegasos-korjauksen onnistumisesta.

Pandemiaseurannan yhteydessä selvisi syksyllä 2020, että COVID-19-testien kvalitatiiviset tulokset oli pääsääntöisesti kirjattu ns. näyttömuodossa rakenteisen kirjaamisen sijasta.

Voisiko sama ilmiö olla totta myös kvantitatiivisten mittareiden kohdalla? Jos näin olisi, muuttuisi laboratoriotulosten poimiminen kertaluokkaa vaikeammaksi. Aineiston pohjalta tekemämme arvio on, että ongelma liittyy sittenkin vain sellaisiin kvalitatiivisiin tutkimustuloksiin kuten jako negatiiviseen ja positiiviseen tulokseen.

Yliopistosairaaloissa käytetyt tietojärjestelmät Uranus ja Esko puuttuvat laboratoriotuloksia kirjaavien järjestelmien listalta. Nämä ovat järjestelmiä, joilla ei ole omaa laboratoriotietojen käyttöliittymää. Uranus-asiakkailta esimerkiksi on erillisjärjestelminä Tamlab (Fimlabin järjestelmä TAYSissa) tai Weblab (Mylabin Multilab-webliittymä). Tiedot kantaan kirjautuvat Uranus-integraatioväylän (Ensemble tai Biztalk) kautta. Laboratorion oma tietojärjestelmä kirjoittaa arkistoon varsinaisen laboratoriotuloksia sisältävän potilasasiakirjan kuvailutietoineen, minkä vuoksi PTA:ssa näiden tulosten tietojärjestelmänimenä on joko Multilab tai LKANTA. Vuonna 2020 Multilab kirjasi 23 % ja Fimlabin LKANTA 7 % kaikista PTA:oon kirja- tuista julkisen terveydenhuollon näyteanalyysistä.

Käytössämme oli pilottiaineisto, joka oli poimittu suoraan Fimlabista sisältäen Pirkanmaan ja Keski-Suomen sairaanhoitopiirin julkisten palveluntuottajien tilaamat laboratorioanalyysit ja niiden tulokset. Tämän aineiston avulla saatoimme selvittää, ovatko Tampereen yliopistolaisen sairaalan näyteanalyysit kirjautuneet em. integraatioväylän kautta potilastiedon arkistoon. Osoittautui, että HbA1c-, LDL-kolesteroli- ja kreatiniini-tutkimusten PTA-kattavuus oli 85–92 % vuosina 2018–2020. Puuttuvaa osaa voivat selittää hienoiset erot esim. näytteenotto- ja analyysipäivien välillä tms. seikat. Asiaa on syytä selvittää vielä tarkemmin.

Fimlab-vertailuaineisto tarjoaisi erinomaisen mahdollisuuden PTA:n laboratoriotietojen validointiin Pirkanmaan ja Keski-Suomen sairaanhoitopiireissä. Valitettavasti emme voineet hyödyntää tätä mahdollisuutta täysimääräisesti, koska tälläkin alueella on laajalti Pegasos käytössä.

Sen sijaan vastaava aineisto Pohjois-Karjalan Siun sotesta osoittaa, että PTA:n laboratoriotulokset voivat parhaimmillaan olla varsin kattavia viime vuosina (ks. [luku 3 Tulokset](#)).

Alueellisissa vertailuissa havaitsimme muutenkin hyvin suurta vaihtelua. Selvitettäväksi vielä jää, löytyykö näiden taustalta muita syitä kuin Pegasos.

Huomiotamme mm. kiinnitti se, että U -AlbKre ja cU-Alb -tutkimusten tulosarvot puuttuivat aineistostamme muita tutkimuksia yleisemmin. Hypoteesimme oli, että mittaustarkkuuden alittavat tai ylittävät arvot ovat prosessoinnin jossain vaiheessa pudonneet pois. Fimlab-aineisto mahdollisti ilmiön tutkimisen parivertailulla: poimimme Fimlab-aineistosta ne tutkimukset, joissa tulokseksi oli merkitty ALLE tai YLI -tyyppinen vastaus. Etsimme vastaavan analyysituloksen PTA-aineistosta potilaan tunnisteeseen, tutkimusnimikkeeseen ja tutkimusajan perusteella. Tämä tarkastelu vahvisti hypoteesin oikeaksi. Asia ei kuitenkaan ollut aivan näin yksinkertainen: myös tarkat numeeriset vastausarvot näyttivät usein kääntyneen puuttuviksi arvoiksi PTA:ssa. Syytä tähän ilmiöön emme ehtineet selvittää, joten se jääkin jatkoprojektin tehtäväksi.

Tutkimme samaa hypoteesia myös Siun sote -pilottiaineistolla. Myös siinä mittaustarkkuuden alittavat ja ylittävät arvot korvautuvat puuttuvalla tiedolla PTA:ssa (esim. Pt-GFRe-MD Yli 60 yhteensä 4089 kpl). Siun soten pilotissa tarkkuuden alittavia tai ylittäviä arvoja oli kuitenkin hyvin vähän, esimerkiksi U-AlbKre-analyysissä ei yhtään – sen sijaan 9500 analyysiä päätyi tulokseen EI LASKETA. Erilaisten käytäntöjen selvittäminen jää myös jatkoprojektin tehtäväksi.

Käytössämme olleet Fimlabin, Siun soten ja Terveystalon pilottiaineistot antoivat erinomaisen mahdollisuuden testata parivertailuna myös yksittäisten laboratoriotulosten löytymistä PTA:sta eri vuosina. Siun soten kohdalla osumistarkkuus oli hyvä vuodesta 2018 lähtien vaihdellen 90 %:n (fP-Kol-LDL) ja 97 %:n (cu-Alb) välillä. Kanta-liittyminen tapahtui vuoden 2017 kuluessa, joten ennen vuotta 2018 vain kolmannes analyysistä löytyi PTA:sta.



Fimlab-tulosten osalta parivertailun tulokset ovat heikot laajan Pegasos-käytön vuoksi: vain 27–29 prosentista (U -AlbKre, B -HbA1c, cU-Alb, fp-Kol-LDL) 41 prosenttiin (P -Krea).

Lisäksi näyttäisi siltä, että albumiinitutkimuksissa käytetään ei-virallisia koodeja yleisemmin, kuin mitä vakiintuneissa HbA1c- ja LDL-analyyseissä. Fimlab-aineistossa esimerkiksi cU-Alb-tutkimuksen virallisen koodin 1029 lisäksi esiintyy usein koodi 35570. Vastaavasti U -AlbKre-tutkimuksen koodina käytettiin 9522 virallisen koodin 4511 sijasta. Aineistomme poiminnassa olivat käytössä vain viralliset Kuntaliiton koodit. Vastedes tulee tutkia em. aggregaattitaulua huolellisesti valittaessa poimittavat laboratoriotutkimukset.

Yksi keskeinen syy hyödyntää Kanta-aineistoa on, että tämä alkaa kattaa myös yksityisen terveydenhuollon. Terveystalon aineisto edustaa laajalti yksityisen työterveyshuollon asiakkaita. Vertailuun soveltui ainoastaan vuosi 2018. Terveystalon tutkimusnimikkeistökäytännöt ovat varsin horjuvat. Vain AlbKre- ja HbA1c-tuloksia löytyi PTA:sta, 34 ja 23 prosenttia vastaavasti. Muut analyysit jäivät löytymättä pääasiassa käytetyn tutkimusnimikkeistön kirjavuuden vuoksi. Tämäkin yksityiskohta vaatii tarkempaa selvitystä ja korjausta jatkossa.

### Tietopyyntöön lisättävät laboratoriotutkimukset

Havaitsimme, että tietopyynnöstä puuttui joitakin Kuntaliiton tutkimuskoodeja, joten tietopyyntöä täydennetään jatkossa puuttuvilta osin.

#### **LDL-kolesteroli**

P-LDL-kolesteroli tulee hakea jatkossa siten, että aineistoon kelpuutetaan plasmasta paastossa ja ei-paastossa otetut LDL-kokeet: fp-LDL (4599), P-LDL (6433).

LDL-kolesterolikoodeja aggregaattitaulusta tutkittaessa paljastui epäilyjä virheellisesti käytetyistä koodeista, kuten 2099, joka on fS-LDL-koodi, mutta tekstimuotoiseen nimikekenttään oli kuitenkin kirjattu "fp-LDL-kolesteroli". Lisäksi 4122 on selkäydinnäytteen koriongonadotropiinin Kuntaliiton koodi, mutta sen selityksenä käytettiin P-LDL-kolesterolia. Näistä virheistä ilmoitetaan rekisterinpitäjille, mutta näitä Kuntaliiton koodeja ei oteta mukaan diabetesrekisteriin.

#### **HbA1c**

HbA1c tulee jatkossa hakea käyttäen kuntakoodeja 6128 (B -HbA1c) ja 6591 (c- HbA1c, "vieri-HbA1c"). HbA1c:n virallinen yksikkö on mmol/mol. Koska HbA1c:tä mitataan toistaiseksi vielä runsaasti erilaisilla vieritesteillä ja jonkin verran käytetään vielä prosentteina ilmoitettavaa HbA1c:tä (1560), josta piti luopua jo 2010, päädyttiin seuraavassa tietohaussa poimimaan tekstihaualla virallisten Kuntaliiton koodien lisäksi kaikki numerosarjat, jotka vastaavat HbA1c-koetta. Jatkossa tuotetaan raportti, kuinka paljon HbA1c-tuloksia löytyy näillä epävirallisilla numeroilla ja pyritään tiedolla vaikuttamaan Kuntaliiton koodien systemaattiseen käyttöön. Prosentteina esiintyvät lukuarvot voidaan tarvittaessa muuntaa kaavalla B -HbA1c (mmol/mol) = 10,93 x B -HbA1c (%) - 23,50.

#### **Kreatiniini ja eGFR**

Aineistopyyntö sisälsi ainoastaan P -Krea-arvot (Kuntaliiton koodi 4600). Tämän lisäksi on syytä poimia fp-Krea (2142), S -Krea (2143) ja fS-Krea.

Pt-GFReEPI on laskennallinen tutkimus, jossa mukana ovat plasman kreatiniinipitoisuus, sukupuoli ja ikä. Potilaan painoa ei tarvita, koska tulos ilmoitetaan normalisoituna aikuisen keskimääräistä kehon pinta-alaa 1,73 m<sup>2</sup> kohti.

eGR-laskemista varten tietopyyntöön otetaan P -Krea (4600), fp-Krea (2142). Lisäksi alkuun otetaan S -Krea (2143) ja fS-Krea, ja jatkossa raportoidaan ne kunnat, joissa sitä vielä käytetään.

### U -AlbKre ja cU-Alb

Tietopyyntöön otetaan 4511 (U -AlbKre), joka on Kuntaliiton kansallinen tutkimuskoodi. Lisäksi poimitaan toistaiseksi myös runsaasti käytetty epävirallinen koodi 9522. Arvoissa esiintyy paljon < tai \*-merkkejä, jolloin arvoksi poimitaan merkkiä seuraava lukuarvo eli mittaus-tarkkuuden alarajan lukema.

Lisäksi poimitaan samaa asiaa mittaava mutta vähitellen käytöstä poistumassa oleva 3557 (cU-Alb) ja lisäksi toistaiseksi sen muunnos 35570.

Viimeisimpää albuminurialukemaa ei käsitellä raporteilla jatkuvana muuttujana vaan U -AlbKrea ja cU-Alb-arvoista tehdään neljä ryhmää: 1) ei mittausta, 2) U -AlbKrea alle 3 tai cU-Alb alle 20, 3) U -AlbKrea 3-30 tai cU-Alb 20-200 sekä 4) U -AlbKrea yli 30 tai cU-Alb yli 200.

### Fysiologiset mittaukset

Verenpaine, pituus ja paino (ja niistä laskettu painoindeksi eli *body mass index*, BMI) olisivat diabetesrekisterin keskeisiä muuttujia. Valitettavasti saamamme aineiston perusteella PTA näyttää sisältävän näitä varsin heikosti rakenteisessa muodossa. Aineistomme sisälsi noin 250000 verenpainemittausta 44000 potilaalta. Vain noin 40000 potilaan pituus- ja painomittauksia oli kirjattu. Näin ollen vain noin 10 prosentilta on tiedossa näitä suureita. Heikon kattavuuden vuoksi emme analysoineet niitä sen tarkemmin.

### Väestötiedot

Käytössämme oli Väestötietojärjestelmästä tiedot henkilöiden sukupuolesta, syntymäajasta, viimeisimmästä kotikunnasta sekä mahdollisesta kuolinpäivästä. Rekisterikopiomme ei kuitenkaan sisällä niitä henkilöitä, jotka olivat kuolleet ennen 8.2.2014. Kokonaisaineistomme sisälsi 8213 sellaista henkilöä, joita ei löytynyt VRK-kopiostamme. Suurimmalta osin nämä henkilöt ovat todennäköisesti kuolleet ennen 8.2.2014. Sen lisäksi aineistomme saattaa sisältää henkilöitä, joilla on väliaikainen henkilötunnus.

Pääsääntöisesti poistimme analyysistä henkilöt, jotka eivät löytyneet väestörekisterikopiostamme.

Hyvinvointialueittain tehdyissä vertailuissa aluejako perustui potilaiden viimeisimpään asuinkuntaan. Joissain kysymyksen asetteluissa tämä aiheuttaa hienoista epätarkkuutta. Jatkossa on harkittava tarkemman muuttohistorian sisällyttämistä henkilöiden tietoihin.

### Lääkitys

Lääkitystiedot perustuivat Reseptikeskuksesta poimituihin lääkemääräys- ja ostotietoihin.

Potilaspopulaation muodostamisessa käytimme lääkemääräystietoja: yksikin diabetesresepti riitti liittämään henkilön osaksi aineistoa. Tämä on mahdollisimman herkkä valinta-ehdo. Aineisto mahdollistaa helposti tiukemmat ehdot, esimerkiksi sen, että lääkettä on myös lunastettu riittävän monta kertaa. Pilottiaineisto antaa hyvän mahdollisuuden erilaisten poimintaehtojen arviointiin. Tämä edellyttää jatkotutkimusta.

Lääkinnälliset alaryhmät muodostettiin potilaan viimeisimmän vuoden aikana ostamien lääkkeiden perusteella. Tässä otettiin huomioon kaikki vuoden sisällä ostetut diabeteslääkkeet, ja alaluokka muodostettiin tämän kokonaisnäkyvän perusteella.

Koko aineistolle poimittiin vielä Reseptikeskuksesta tiedot muusta lääkityksestä kuten verenpaine- ja lipidilääkitys.

Reseptikeskuksen käyttäminen tietolähteenä voi jättää ulkopuolelle esim. laitoshoidossa olevia potilaita, joille ei välttämättä laadita reseptejä.

Sähköinen resepti otettiin käyttöön vuonna 2010, ja pakollinen se on ollut vuodesta 2017 lähtien.

Arvioimme hoitomyöntyvyyttä (lääkeadherenssia) varsin karkeasti selvittäen, kuinka suuri osuus potilaista oli viimeisten 120 päivän aikana lunastanut hänelle määrättyjä diabeteslääkkeitä. Näiden potilaiden osuus oli 81 %, mitä voitaneen pitää kohtalaisen hyvänä. Jatkossa hoitomyöntyvyyttä on tärkeää tarkastella etenkin potilasryhmissä, joissa havaitaan HbA1c- ja LDL-tulosten perustella huono hoitotasapaino. Näin saadaan tärkeää tietoa yleisimmistä hoitotasapainon taustalla olevista syistä, eli johtuuko huono hoitotasapaino todennäköisimmin puuttuvasta lääkemääräyksestä, puuttuvasta lääkeadherenssista, liian vähäisestä annoksesta vai väärän tyyppisestä lääkityksestä.

Tehtäväksi jää myös laskea tarkemmat tunnusluvut lääkeadherenssille.

## Johtopäätökset

Käyttämämme tietolähteet tarjoavat nähdäksemme kohtalaisen perustan kansalliselle diabetesrekisterille. Keskeiset rekisterin tietoelementit ovat saatavissa elämäntapamuuttujia (kuten tupakointi, liikunta), fysiologisia mittauksia (verenpaine, lihavuus) ja PROM- eli potilaiden itsensä raportoimia vointitietoja lukuun ottamatta. Lukuisat työn kuluessa havaitsemamme mallinnus- tai poimintavirheet sekä -puutteet on saatu jo pilotin kuluessa korjattua. Näinä aikoina myös potilastietojärjestelmät joutuvat uusimaan ohjelmansa vastaamaan uusia, ns. 2016 Kanta-määrittämissä. Nykyisin käytössä olevat Kelan yhteistestauskäytännöt ovat aivan eri tasolla kuin viisi vuotta sitten, mikä antaa odottaa myös entistä yhtenäisempiä tallennuskäytäntöjä tulevaisuudessa.

Diabetesta ja sen komplikaatioita hoidetaan laajalti niin yksityisessä kuin julkisessa terveydenhuollossa, erikoissairaanhoidossa ja perusterveydenhuollossa. Vastaavan aineiston kerääminen näistä kaikista eri lähteistä jollain toisella tavalla tuskin olisi kustannustehokasta.

Helposti analysoitavan, relaatiotietokantamuotoisen pilottirekisteriaineiston rakentaminen potilastiedon hierarkkisesta asiakirja-arkistosta on ollut suuri ponnistus, joka toimii pioneeri-esimerkkinä muillekin PTA:n toisiokäyttöä harkitseville hankkeille. Erytyskiitos tämän työn tekemisestä lankeaa Kelan Tietoallas-tiimille. Ponnistuksen ei soisi valuvan hukkaan.

Jäljellä on toki vielä paljon tarkistettavaa ja seurattavaa, mikä kuitenkin on ominaista kaikelle rekisterityölle. On mahdollista, että myös kirjauskäytänteiden ohjeistusta pitää tarkentaa ja ohjeiden noudattamista korostaa. Toisaalta, kun rekisteri tuottaa raporttiaineistoa esimerkiksi vertaisoppimisen ja -kehittämisen tueksi, on jo sillä itsellään ohjaava vaikutus.

## 5 Jatkosuunnitelma ja suositukset

PTA-aineistoon, lääkekorvattavuustietoihin ja eResepti-tietoihin perustuva kansallinen diabetesrekisteri todettiin pääosin toimivaksi. Kehitys- ja korjaustarpeita havaittiin, ja kehittämistyö jatkuu Valtava-hankkeen toisella hankekaudella Laboratoriotietojen Kanta-tietotutkimusprojektissa sekä THL:n Laaturekisteriprojektissa. Tavoitteena on valtakunnalliseksi skaalautuva PTA:n diagnooseihin ja laboratoriokokeisiin, Reseptikeskuksen lääkemääräys- ja ostotietoihin, Kelan lääkekorvausoikeuksiin sekä DVV:n väestörekisteritietoihin perustuva diabetesrekisteri, jonka pohjalta voidaan käynnistää kansallinen vertaiskehittäminen.

Diabetesrekisterissä ja laajemmin myös muissa laaturekistereissä on tunnistettu tarve tietoihin, joita ei toistaiseksi tallenneta Kanta- eikä muihin kansallisiin tietovarantoihin rakenteisessa muodossa. Tällaisia tietoja ovat esimerkiksi fysiologiset mittaukset, tupakointitieto, kliiniset löydökset, sairaalalääkitykset, verensiirrot ja elämänlaatu- sekä muu vointitieto. Kanta-tietovarantojen laajempi käyttäminen laaturekisterien tiedonlähteinä edellyttää siis myös huomattavaa rakenteisen Kanta-tietosisältöjen kehittämistä.

Projektin aikana havaitut korjausehdotukset kirjataan Valtava-hankkeen jatkotoimintasuunnitelmaan ja uuteen tietopyyntöön, jota edelleen tarkennetaan Kelan ja THL:n asiantuntijoiden ja kliinikkoasiantuntijoiden vuoropuhelussa Q1/2022.

Potilaspopulaatio muodostetaan jatkossa seuraavilla määrittelyillä

- Potilastiedon arkistosta (PTA) potilaat, joilla on potilaan elinaikana esiintynyt pää- tai sivudiagnoosina diabetesdiagnoosi ilman aikarajausta (ICD-10: E10\*, E11\* tai E12\*–E14\*)
- Kelan erityiskorvausoikeusrekisteristä (Etuus) potilaat, joille on potilaan elinaikana myönnetty diabeteslääkkeiden erityiskorvattavuusoikeus 103 tai 215 sekä 250, 285, 295, 346, 3011. Tältä osin on huomattava, että vuonna 2016 tapahtuneen perus- ja erityiskorvattavuusoikeuden erottamisen myötä kaikille aiemmin 103-oikeuden saaneille kirjattiin myös 215-oikeus.
- Lääkitystietokannasta (E-Resepti) potilaat, joille on potilaan elinaikana määrätty diabeteslääkkeitä (ATC-koodi A10\*)

Potilaiden tunnistamiseen ja diabetestyyppin arvioimiseen tarvittavien diagnoosien kattavuus tulisi saada kaikkien hyvinvointialueiden osalta arviolta 80 %:iin. Diagnoosien perusteella tunnistettujen potilaiden osuus summapopulaatiosta arvioidaan ja raportoidaan. Pyritään tunnistamaan teknisiä ongelmia niillä hyvinvointialueilla, joissa diabetesdiagnoosien kirjaimiskattavuus jää alle tavoitteen.

Pyritään validoimaan iskeemisten sydänsairauksien ja loppuvaiheen munuaisten vajaatoiminnan PTA-diagnooseihin perustuvaa esiintyvyyttä ja ilmaantuvuutta viimeisten 2 vuoden aikana.

Keskeisin ongelma oli laboratoriokokeiden havaittu systeemivirhe Pegasoksen laboratoriotutkimuskoodeiden siirtymisessä PTA:oon, mikä korjautuu todennäköisesti Q1/2022. Tämän jälkeen pyydetään uusia aggregoidut laboratoriotutkimuskoodeja on käytetty Kuntaliiton koodien lisäksi HbA1c-, LDL-kolesteroli-, krea- ja U-AlbKre-tutkimuksissa. Jotta saadaan myös historialliset laboratoriotutkimukset mukaan raportteihin, käytetään myös koko historiatiedon kattavaa PTA:oon kirjattujen laboratoriotutkimusten aggregoitua aineistoa ("aggregoidut\_labrat").

Laboratorioita, sote-rekisterinpitäjiä ja tietojärjestelmätoimittajia ohjeistetaan kiinnittämään huomiota kansallisen laboratorionimikkeistön käyttöön. Tähän velvoittaa myös [THL:n määräys sosiaali- ja terveydenhuollon tiedonhallintaan 5/2021 \(liite 2, välilehti Toiminnot, VV16\)](#):

- Asiakastietoja tai -asiakirjoja kansallisiin tietojärjestelmäpalveluihin lähettävän järjestelmän ja tiedot tuottavan palvelunantajan on tuotettava lähetettävien asiakirjojen tai sanomien koodisto-tyyppisiin tietokenttiin sisältö noudattaen THL:n koodistopalvelun kautta käsiteltyjä ja julkaistuja koodistoja ja luokituksia, ellei tarkemmassa tietosisältömäärittelyssä muuta ilmaista.
- Vaikka vastaaviin tietoihin käytettäisiin paikallisesti muita tai täydennettyjä koodistoja, on Kanta-palveluihin lähetettävissä tiedoissa tieto ilmaistava kansallisen koodiston mukaisesti ja mahdollisten vastaavuuksien määrittely on lähettävän toimijan vastuulla. Paikallisten koodistojen käyttö voi olla mahdollista ja tarpeellista joissakin käyttökohteissa, jolloin kansallisiin tietojärjestelmäpalveluihin lähetettäviin kansallisten koodistojen mukaisesti tietoihin on mahdollista määrittellä vastaavuudet. Jos samoja koodistosisältöjä julkaistaan useissa eri kanavissa, on noudatettava koodistoversio ja sen käyttöönotto-ajankohta ensisijaisesti THL:n kautta julkaistu. Sisältö- tai luokituskohtaisesti on mahdollista poiketa perusvaatimuksesta THL:n tarkempien ohjeiden ja määrittelyjen mukaisesti.

Diabetesrekisterin uutena muuttujana määritetään myös diabeteksen kesto. Potilaspopulaation muodostamisen yhteydessä on kerättävä tieto varhaisimmasta ajankohdasta, kun potilaalla esiintyy jonkun kolmen tietolähteen perustella diabetes: Etuusrekisteri ja lääkekorvattavuuden myöntövuosi, eResepti ja ensimmäisen diabeteslääkkeen reseptin laatimisen vuosi sekä PTA ja ensimmäisen diabetesdiagnoosin esiintymisvuosi. Kesto lasketaan vähentämällä varhaisin diabeteksen esiintymisvuosi tietopointivuodesta.

Kuolleisuutta ja keskeisten diabeteksen komplikaatioiden esiintymistä arvioidaan suhteessa diabeteksen keston, seuranta-aikaan sekä erikseen alaryhmissä (ikä, sukupuoli, diabetes-tyyppi, lääkeryhmät). Diabeteksen hoitotasapainoa kuvaavat indikaattorit esitetään tarkennetuissa hoidollisissa alaryhmissä (oheissairastavuus ja riskiluokitus).

Uudessa aineistossa tehdään pääosin samat analyysit, kuin nyt päättyvässä loppuraportissa sekä 2018–2020 pilotissa.

Tulokset esitetään niille hyvinvointialueille, joilla diagnoosi- ja laboratoriokoekattavuus on yli 80 prosenttia. Mikäli valtaosassa hyvinvointialueita saadaan luotettavaa laatutietoa, käynnistetään raporttien perusteella kansallinen vertaiskehittäminen hyvinvointialueittain.

# 6 Teknologialiite

## Tausta

Tämä materiaali on Diabetesrekisterin tietoaalaspilotti -projektin loppuraportin teknologia-liite. Vuosina 2020–2021 toteutettu projekti oli osa [STM:n Toivo-ohjelman Valtava-hanketta](#). Projekti toteutettiin yhteistyössä Valtava-hankkeen Kelan projektin "Potilastiedon arkisto tietoaalaseen (PTA)" sekä THL:n laaturekisterien kehittämisen projektin kanssa.

## Potilastiedon arkisto PTA

### Asiakirjan rakenne

Potilastiedon arkisto (PTA) koostuu hierarkkisista XML-dokumenteista, joita kutsutaan *asiakirjoiksi* tai *CDA-dokumenteiksi*. Akronyymi CDA viittaa [HL7-organisaation](#) määrittelemään XML-rakenteeseen (*Common Data Architecture*). Tärkeimmät *asiakirjatyypit* ovat *hoitoasiakirja* ja *palvelutapahtuma-asiakirja*.

Jokaisella asiakirjalla on yksikäsitteinen OID-tunnisteensa. Kun asiakirjaan tehdään muutoksia tai lisäyksiä, luodaan uusi asiakirja. Aikaisemmat asiakirjan versiot yhdistetään toisiinsa ensimmäisen asiakirjaversioiden tunnuksella, jota kutsutaan nimellä *setId* (CDA-skeemassa) tai *AlkuperäinenAsiakirjaOID* (05-tason mallinnuksessa).

**Palvelutapahtuma** sitoo yhteen useita hoitoasiakirjoja. Yhdistävä lenkki on ns. palvelutapahtuman OID ("palvelutapahtumatunnus").

Tärkein asiakirjatyypin on hoitoasiakirja (*Clinical Document*), ks. [XML-skeema](#). Sen juuritasolla on kerrottu mm. asiakirjan OID-tunniste (*id*), alkuperäisen asiakirjan tunniste (*setId*) ja rekisterinpitäjän tunniste (*custodian*). Lisäksi juuritasolta löytyy kansallinen laajennuselementti *hl7fi:localHeader*, joka sisältää mm. asiakirjan kirjanneen tietojärjestelmän nimen ja ohjelmaversioiden sekä ohjelman valmistajan tunnistet.

Varsinainen hoitoasiakirjan sisältö on kirjattu *component*-nimiseen elementtiin. Se koostuu joko rakenteisesta asiakirjarungosta (*structuredBody*) tai vaihtoehtoisesti ei-rakenteisesta sisällöstä *nonXMLBody*.

Rakenteinen runko koostuu kolmesta sisäkkäisestä *component/section*-elementistä. Näitä nimitetään *näkymäksi*, *hoitoprosessin vaiheeksi* ja *otsikoksi* ulommasta sisempään lueteltuina. Niiden keskinäiset suhteet ovat yksi moneen, ts. näkymä voi sisältää yksi tai useampia hoitoprosessin vaiheita, joista kukin voi sisältää yhden tai useamman alaluvun eli otsikon. Ylintä tasoa nimitetään myös usein *merkinnäksi*. Asiakirja voi siis sisältää useita merkintöjä. Paperisten asiakirjojen aikaan nämä merkinnät oli kirjattu usein erivärisille papereille riippuen siitä, oliko kyse sisätauteihin, radiologiaan, kirurgiaan, diagnosointiin tms. liittyvä näkymä (ks. [koodistopalvelu AR/YDIN - Näkymät, OID=1.2.246.537.6.12.2002](#)). Alin, eli otsikkotasoa voi sisältää useampia *entry*-elementtejä, joiden sisällöstä löytyvät varsinaiset rakenteiset tulokset kuten diagnoosit, fysiologiset mittaukset, laboratoriotulokset (*observation*-elementti) tai toimenpiteet (*procedure*-elementti).

Entry-elementti ei ole pakollinen, ts. otsikkotasoa ei välttämättä sisällä lainkaan tuloksia em. muodossa. Sen sijaan tulokset on voitu kirjata pelkästään ns. *näyttömuodossa text*-nimisen otsikkotason elementin sisältönä. Kyseinen elementti noudattaa HL7-organisaation määrittelemää [StrucDoc-määrittelyn Text-elementin](#) rakennetta, joka muistuttaa paljon HTML-

merkintäkieltä, tärkeimpinä juurielementteinään *table* ja *paragraph*. Useimmat ohjelmistot kirjaavat edelleen esimerkiksi rokotustiedot pelkästään näyttömuodossa. Jotkut ohjelmistot kirjaavat myös laboratoriotietoja ainoastaan näyttömuodossa. Tämä vaikeuttaa oleellisesti ao. tulosten toisiokäyttöä.

Potilastiedon arkisto (PTA) koostuu kaiken kaikkiaan

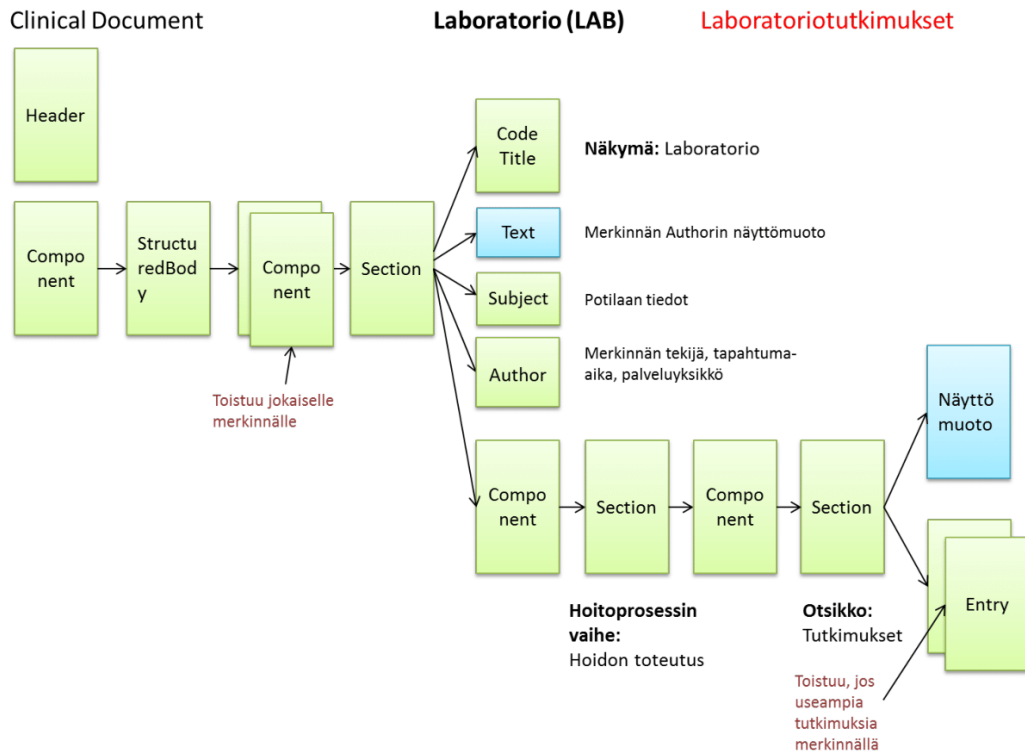
1. HL7 CDA R2 -rakenteisista XML-asiakirjoista (Documentum-tietokanta)
  - a. hoitoasiakirjat
  - b. palvelutapahtumat
  - c. suostumukset yms. lomakeasiakirjat
2. kuvailutiedoista ("metatiedoista")
3. lokitiedoista.

Hierarkkinen XML-asiakirjojen tekstitietokanta soveltuu huonosti toisiokäyttöön ja potilaan tietojen yhteenvetojen esittämiseen ("koosteisiin").

### CDA R2 hoitoasiakirjan rakenne



Esimerkkinä laboratoriotutkimukset ja niiden tulokset:

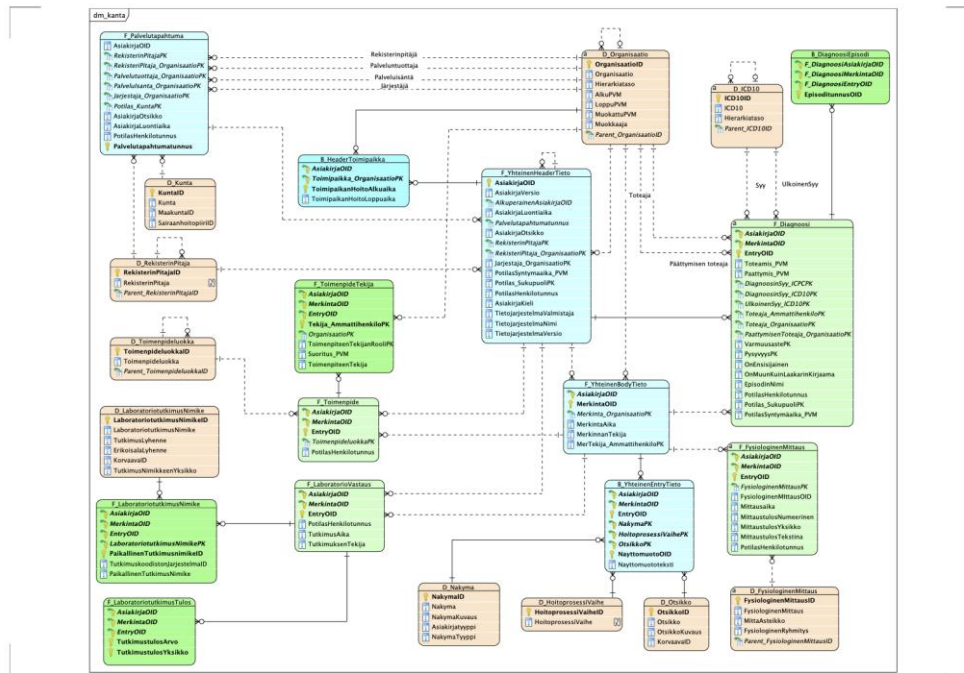


Vaiheen 5 mallinnus

## Asiakirjat ja niiden versiot

Kelassa on toiminut vuodesta 2018 lähtien tiimi, joka on kehittänyt PTA:n hierarkkisesta mallista ns. tähtimallisen relaatiotietokantaskeeman. Päämallintajana on toiminut Jyrki Kyläheiko. Hierarkkisen rakenteen purkaminen ko. malliin on monivaiheinen prosessi. Valmistusta tähtimallia on totuttu kutsumaan vaiheen 5 tasoksi. Se on toiminut PTA-dm-pilotin analyysikannan lähdeaineistona. Rekisterin kannalta keskeisimmät faktat ja dimensiot on kuvattu seuraavassa ER-mallissa:





ER-kaavio on vain viitteellinen: sen fakta- ja dimensiotauluihin on poimittu sarakkeita valikoitujen. Lisäksi on huomattava muutama historiallinen seikka:

1. Talvella 2021 siirryttiin dimensioviittauksissa surrogaattiavaimista ("PK") ns. luonnollisiin avaimiin. Tämä tarkoittaa sitä, että faktataulujen (F-alkuiset) PK-loppuiset sarakkeet eivät tosiasiallisesti olekaan generoituja surrogaattiavaimia, vaan dimensiotaulujen todellisia, ID-loppuisia arvoja (jos ko. arvoja nyt ylimalkaan dimensiotaulusta löytyykään). Tällä epäortodoksisella muutoksella haluttiin välttää tilanteet, joissa viittaukset osoittivat pelkkään "Tuntematon"-arvoon.
2. Potilaan kryptattu tunniste (*PotilasHenkilotunnus*) sekä syntymäaika ja sukupuoli olivat aluksi toistettu jokaiseen faktaan erikseen. Sitten keväällä 2021 ne siirrettiin tauluun *F\_YhteinenHeaderTieto*. Nyt ao. sarakkeet esiintyvät useassa paikassa.
3. Useimpien sarakkeiden Oracle-tietotyyppi on NVARCHAR2, mikä pitää ottaa huomioon merkkijonovertailuissa yms.

Mallinnuksessa kaikille asiakirjoille yhteiset attribuutit on koottu faktatauluun *F\_YhteinenHeaderTieto*. On hyvä huomata, että ao. taulu sisältää rivin kustakin asiakirjan versiosta. Tämän vuoksi analyyseissä on tyypillisesti pitäydyttävä kunkin asiakirjajoukon viimeisimpään versioon. Tämä valitaan kenttien *AlkuperainenAsiakirjaOID* (vastaa XML-elementtiä *setId*) ja *AsiakirjaVersio* avulla (huomaa että jälkimmäisen tietotyyppi on merkkijono jonka lajittelujärjestys eroaa numeerisesta arvosta!).

Rekisterinpitäjyys on sikäli kaksijakoinen, että julkiset rekisterinpitäjät on kirjattu omaan rekisterinpitäjärekisteriinsä (Koodistopalvelussa [eArkisto - Rekisterinpitäjärekisteri 2011](#)). Tämä koodisto on kopioitu dimensioon *D\_RekisterinPitaja*. Sen sijaan yksityiset rekisterinpitäjät eivät ole tässä joukossa, vaan nämä tunnistetaan organisaatiotunnuksensa perusteella (taulun *F\_YhteinenHeaderTieto* sarakkeella *RekisteriPitaja\_OrganisaatioPK*).

## Palvelutapahtumat

Edellä olevassa ER-kaaviossa on taulun *F\_Palvelutapahtuma* primaariavaimiksi merkitty harhaanjohtavasti sarakkeella *Palvelutapahtumatunnus*. Tosiasiassa avain on tietysti *AsiakirjaOID*,

mutta koska palvelutapahtumien asiakirjatunnuksilla ei ole mitään tekemistä hoitoasiakirjojen tunnuksien kanssa, on kuvaan tehty tämä vääristymä, jotta voitaisiin jotenkin kuvata palvelutapahtumien ja hoitotapahtumien (*F\_YhteinenHeaderTieto*) välinen yhteys. Tosiasiassa siis samalla palvelutapahtumatunnuksella on olemassa monta riviä faktataulussa *F\_Palvelutapahtuma*.

Palvelutapahtuma on jossain määrin diffuusi käsite. On olemassa tilanteita, joissa täsmälleen sama yksittäinen hoitokäynti on kirjattu kahdeksi eri palvelutapahtumaksi. Toisaalta taas ei ole takeita myöskään siitä, että kaikki tiettyyn palvelukokonaisuuteen kuuluvat hoitoasiakirjat olisi kirjattu samalle palvelutapahtumalle. On myös niin, että yksittäiseen palvelutapahtumaan on voitu kirjata hoitoasiakirjoja useiden vuosien ajalta – ehkä syystäkin.

## Merkinnot

Yksittäinen hoitoasiakirja voi sisältää yhden tai useamman merkinnän. Merkinnän primariavain koostuu sarakkeista *AsiakirjaOID* ja *MerkintaOID*. Näin muodoin taulun *F\_YhteinenBodyTieto* pitäisi sisältää rivin kustakin asiakirjamerkinnästä.

## Entryt

Kaikki 05-tasolla olevat faktarivit on poimittu sillä ehdolla, että asiakirjasta löytyy *entry*-elementti *otsikko*-tason alta. Kuten yllä on todettu, on mahdollista, ettei rakenteinen asiakirja (so. sellainen, jossa on elementti *structuredBody*) kuitenkaan sisällä *entry*-elementtiä, vaan ainoastaan ns. näyttömuodon *otsikko*-tason alla olevassa *text*-elementissä. Tämä on tärkeä pitää mielessä.

Taulu *B\_YhteinenEntryTieto* tuli mallinnuksen osaksi suhteellisen myöhään, vasta keväällä 2021. Siinä pyrittiin ensinnäkin korjaamaan aikaisemmassa mallinnuksessa ollut virhe, jossa oletettiin, että merkintä ei voisi sisältää kuin yhden otsikkotason elementin. Nyt tämä taulu sisältää listan asiakirjamerkinnän kaikista hoitoprosessin vaiheista ja näiden otsikoista. Lisäksi on viite *entry*-elementtiin (*EntryOID*). Mallinnuksessa on huomioitu myös se mahdollisuus, että *entry*-elementtiä ei ole, vaan ainoastaan näyttömuoto (*NayttomuotoOID*). Tämä on tässä yhteydessä lähinnä teoreettinen mahdollisuus, koska poiminnat tehdään *entry*-elementin sisällön perusteella. Tauluun on kuitenkin lisätty sarake *Nayttomuototeksti*, joka sisältää siis *StrucDoc.Text*-muotoisen rakenteisen näyttömuodon (tosin tätä kirjoitettaessa JSON-muotoon konvertoituna).

## Diagnoosit

PTA:ssa esiintyvät diagnoosit on poimittu faktatauluun *F\_Diagnoosi*. Taulussa on ICPC2-luokituksen mukainen diagnoosi ja sen lisäksi peräti viisi ICD10-diagnoosikoodia, jotka voivat olla ns. yhdistelmäkoodeja. Käytännössä vain *DiagnoosinSyy\_ICD10PK* ja *UlkoinenSyy\_ICD10PK* sisältävät merkityksellisiä arvoja. Sarake *OnEnsisijainen* sisältää periaatteessa tiedon siitä, onko kyseessä päädiagnoosi vai ei. Sarakkeen *Toteamis\_PVM* tulisi kai periaatteessa sisältää tiedon siitä ajankohdasta, jolloin esim. diabetesdiagnoosi on ensimmäisen kerran asetettu. Näin ei yleensä ole, vaan kyseessä on enemmänkin lääkärin tapaamisajankohta (vastaa aikaleimaa *AsiakirjaLuontiaika* taulussa *F\_YhteinenHeaderTieto*). Diagnoosin päättymispäivämääriä on mitä ilmeisimmin vähän kirjattu.

## Toimenpiteet

Kirurgiset yms. toimenpiteet on poimittu tauluun *F\_Toimenpide*. Toimenpiteellä voi olla monta tekijää ja se voi jatkua usealle päivälle. Sen vuoksi on olemassa taulu *F\_ToimenpideTekija*, joka paitsi kertoo toimenpiteen tekijät ja suorituspäivämäärät myös siltaa toimenpiteen suorittaviin organisaatioihin.

## Laboratoriotulokset

Jokaisesta laboratoriovastauksesta on rivi faktataulussa *F\_LaboratorioVastaus*. Laboratoriovastaus voi mallinnuksen perusteella sisältää usean eri nimikkeen tuloksen. Tämän vuoksi vastaukseen liittyvät analyysit on kirjattu omaan tauluunsa *F\_LaboratoriotutkimusNimike*. Nimike tulisi määritellä Kuntaliiton tutkimusnimikkeistöllä (dimensio *D\_LaboratoriotutkimusNimike*), mutta käytännössä monasti käytetään paikallista tutkimusnimikkeistöä. Tämä tunnus kirjataan ohjeiden mukaan sarakkeeseen *PaikallinenTutkimusnimikeID*. Asiakirjojen rakenteiden eroavuuden vuoksi joidenkin järjestelmien kirjaukset päätyvät tähän kenttään, vaikka sisältävätkin virallisen Kuntaliiton koodin. Paikallisen tutkimusnimikkeen mnemonisen tunnisteen tulisi löytyä sarakkeesta *PaikallinenTutkimusNimike*. Kanta-DM-pilotissa kaikki laboratoriotulokset poimittiin pelkästään virallisilla Kuntaliiton tutkimuskoodeilla.

Samasta laboratorioanalyysistä voi periaatteessa olla tuloksena useita eri arvoja. Lähinnä tämä liittyy siihen, että arvo voidaan ilmoittaa eri mittayksiköissä. Näin ollen analyysien nimikkeet ja tulosarvot on tallennettu eri tauluihin, tulokset tauluun *F\_laboratoriotutkimusTulos*. Malli tietysti mahdollistaa myös sellaisen virhetilanteen, jossa analyysistä on olemassa kaksi eri tulosarvoa samallakin mittayksiköllä.

## Fysiologiset mittaukset

Fysiologiset mittaukset on kirjattu faktatauluun *F\_FysiologinenMittaus*. Aineistopyynnön mukaisesti poimittiin seuraavat mittaustulokset:

```
create view dm_findingpk as
select fysiologinenmittausid
from d_fysiologinenmittaus
where fysiologinenryhmitys like 'Kehon mitat%'
      or fysiologinenryhmitys like '%/Verenpaine';
```

[Aineisto ja menetelmät -luvussa](#) on tarkka listaus ehdon täyttävistä mittauksista.

## Analyysitietokanta

### Nimeämiskäytänteistä

Taulujen ja sarakkeiden nimet ovat englanninkielisiä silloin, kun niillä ei ole suoraa vastinetta alkuperäisessä tähtimallisessa Kanta-skeemassa. Suomenkieliset nimet pyrkivät olemaan täsmälleen samoja kuin Kanta-vastineensa.

Taulujen etuliitteillä on pyritty ryhmittelemään tauluja käyttötarkoituksensa ja EARP-elinkaarimallin mukaan:

- f\_ viittaa alkuperäisiin Kanta-faktatauluihin
- b\_ viittaa alkuperäisiin Kanta-siltatauluihin (bridge)
- d\_ viittaa dimensiotauluihin, jotka voivat olla alkuperäisiä Kanta-dimensioita tai analyysikantaa varten luotuja
- dm\_ viittaa diabetesrekisterin muihin metatietotauluihin
- e\_ viittaa ns. entry-tason tauluihin (Earp) – tässä skeemassa näitä ei juuri esiinny, koska alkuperäiset f\_ ja b\_-taulut oikeastaan edustavat niitä
- a\_ viittaa analyysitauluihin (eArp)
- v\_ viittaa näkymiin, jotka usein ovat analyysitaulujen määritelmiä tai sitten analyysitauluihin rinnastettavia entiteettejä
- r\_ viittaa raportointitason tauluihin, jotka sisältävät yleensä jo aggregoitua dataa (eaRp); ovat esimerkiksi tilastokuutioita.

Esimerkiksi taulu a\_diagnoosi on analyysitaulu, jolle on käsitteellinen vastine Kanta-tietokannassa (f\_diagnoosi). Tämän vuoksi taulun nimi on suomeksi ja on yhteneväinen Kanta-nimeämisen kanssa. Toisaalta a\_patient on analyysitaulu, jolle ei ole suoraa vastinetta Kanta-skeemassa. Sen vuoksi taulun nimi on englanninkielinen.

Kanta-skeeman nimeämiskäytänteitä ovat mm.

- dimensioiden primaariavaimet päättyvät merkkijonoon "pk" (primary key); faktataulujen dimensioiden viittaavat vierasavaimet on nimetty vastaavalla tavalla;
- dimensioiden esiintyvät koodiarvot päättyvät "id"-loppuun
- dimensioiden selitysteksti on samanniminen kuin dimensiotaulun nimikin
- oid-tyyppisten sarakkeiden nimet päättyvät "oid", esimerkiksi "asiakirjaoid", "merkintaoid", "entryoid" ja "palvelutapahtumaoid".

Esimerkiksi dimensiotaulun d\_icd10 primaariavain on "icd10pk", icd-koodiarvon sarake "icd10id" ja diagnoosin selitysteksti "icd10".

## Diagnoosit

### Lähdeaineisto F\_Diagnoosi

Kanta-DM-pilotissa diagnoosien faktatauluun *F\_Diagnoosi* Kela poimi kaikki sellaiset diagnoosifaktat, joissa mikä tahansa diagnoseista löytyy näkymästä dm\_icd10id:

```
create view dm_icd10id as
select icd10id from d_icd10
where
-- diabetes
  regexp_like( icd10id , 'E1[01234]|O24.4' )
-- infektiot
  or regexp_like( icd10id,

' [AB]\d{2}|G0\d|H60.[0123]|H6[68]| [IJ]02|J8[56]|K6[567]|K35|L0[0123
45678]|M0[0123]|N[13]0|N39|H60.9|K75.0|T81.4')
-- silmäsairaudet
  or regexp_like( icd10id,
'H36.[012345679]|H43.1|H45.0|H54|H40.5|H42.0')
-- neuropatia
  or regexp_like( icd10id,
'M14.[26]|G59.0|G63.2|G73.0|G99.0|N48.4')
-- jalkaongelmat
  or regexp_like( icd10id,
'L97|I7[09].2|L28|L85.3|L84|M86.6|M72.2|L60.0|M21.[45]|M20.[124]|Z8
9.[4567]')
-- munuaissairaus
  or regexp_like( icd10id,
'N08.3|N18|Z49|Z94.0|I2[012345]|I46|I50|I48|I63.[012345789]|I6[456]
|I7[09].2')
-- lihavuus, uniapnea, rasvamaksa tai NASH
  or regexp_like( icd10id, 'E66|G47.3|K76.0|K73.80')
-- suun terveys, mielenterveys, alkoholi, huumeet, endokrinologiset
sairaudet
  or regexp_like( icd10id, 'K0[45]' -- suun terveys
  || 'F3[23].[123]|F34.1|F0[0123]|F1[01].73|F20|F31|F[17]\d|F50'
-- mielenterveys alkoholi huume
  ||
'E0[01234567]|E16|E2\d|E3[012345]|K5[012]|K73|K74.3|K75.4|K90.0|M0[
56789]|M1[01234]|M8[01]')
;
```

On huomattava, että hakuehtoa on tarkasteltava jatkossa kriittisesti. Esimerkiksi alkoholi- ja huumesairaudet (rivi 21) eivät tulleet tällä haulla poimittua, koska hakuehdon `F[17]\d` olisi tullut olla muotoa `F[17]\d+`.

### Analyysitaulu A\_Diagnoosi

Näkymä *V\_Diagnoosi* ja siitä edelleen analyysitaulu *A\_Diagnoosi* luodaan tästä faktataulusta seuraavasti.

```
create or replace view v_diagnoosi as
with latest_version as (
  select * from (
    select hextoraw( potilashenkilotunnus ) as pid,
      asiakirjaoid, alkuperainenasiakirjaoid,
      asiakirjaluontiaika, rekisterinpitajapk, rekisteripitaja_organisaatiopk,
      palvelutapahtumatunnus, asiakirjaversio,
      row_number() over ( partition by alkuperainenasiakirjaoid order
by to_number(asiakirjaversio) desc ) as seq
    from f_diagnoosi inner join f_yhteinenheadertieto using (asiakirjaoid)
```

```

)
where seq=1
)
select sys_guid() as dgid,
       hexoraw( potilashenkilotunnus ) as pid,
       case when regexp_like( toteamis_pvm, '\d{8}' ) then to_date(
substr( toteamis_pvm, 1, 8), 'YYYYMMDD') else trunc(asiakirjaluontiaika) end as dgdate,
       case when regexp_like( paattymis_pvm, '\d{8}' ) then to_date(
substr( paattymis_pvm, 1, 8), 'YYYYMMDD') else null end as
enddate,
       case when regexp_like( potilassyntyma aika_pvm, '\d{8}' ) then
to_date( potilassyntyma aika_pvm, 'YYYYMMDD') else null end as
bdate,
       decode( potilas_sukupuolipk, '1', 'mies', '2', 'nainen', '3',
'muu', 'tuntematon' ) as gender,
       diagnoosinsyy_icd10pk as syy_icd10id,
       ulkoinensyy_icd10pk as ulksyy_icd10id,
       toteaja_organisaatiopk,
       asiakirjaoid, merkintaoid, entryoid,
       alkuperainenasiakirjaoid, asiakirjaluontiaika, palvelutapahtuma-
tunnus, rekisterinpitajapk, rekisteripitaja_organisaatiopk
from latest_version inner join f_diagnoosi d using (asiakirjaoid)
where paattymis_pvm <= 0;

create table a_diagnoosi as select * from v_diagnoosi;
alter table a_diagnoosi add constraint a_diagnoosi_pk primary key (
dgid );

```

Analyyssitaulua luotaessa täytyy siis pitäytyä kunkin diagnoosin viimeisimmässä merkintäversiossa (aliksely *latest\_version*).

### Palvelutapahtuman tunnistaminen

Seuraavalla ohjelmapätkällä yritetään liittää viimeisimpään diagnoosimerkintään siihen kuuluva palvelutapahtuma (sarake `pt_asiakirjaoid`) käyttäen palvelutapahtumatunnusta sekä palvelutapahtumamerkinnän ja asiakirjamerkinnän luontiaikaa yhdistävinä tekijöinä:

```

alter table a_diagnoosi add ( pt_asiakirjaoid varchar2(100));
merge into a_diagnoosi d using (
select * from (
select dgid, pt.asiakirjaoid,
       row_number() over (partition by palvelutapahtumatunnus
order by abs( extract( day from (h.asiakirjaluontiaika
- pt.asiakirjaluontiaika )*24*60*60))) seq
from a_diagnoosi p
inner join f_yhteinenheadertieto h on p.asiakirjaoid =
h.asiakirjaoid
inner join f_palvelutapahtuma pt using (palvelutapahtumatun-
nus)
) where seq = 1
) s
on ( d.dgid = s.dgid )
when matched then update set d.pt_asiakirjaoid = s.asiakirjaoid;

```

## Potilaiden identifiointi

### Nojautuen PTA:n ICD10-diagnosiin

Diabetespotilaiden taulun *A\_Patient* on määrä sisältää unioni kaikista niistä potilaista, joiden arvellaan jonkin tietolähteen perusteella sairastavan diabetestä. Taulun rakentaminen aloitetaan poimimalla edellä luodusta *A\_Diagnosi*-taulusta kaikki yhdenkin diabetesdiagnoosin omaavat henkilöt. Tämä tapahtuu proseduurin *update\_patient()* avulla:

```
create or replace procedure update_patient as
begin
insert into a_patient (src, dgid, pid, asiakirjaoid, syy_icd10id,
ulksyy_icd10id, dgdate,
  bdate, gender, dmttype, rekisterinpitajaid, rekisteripitaja_organisaatioid )
select dgid, pid, asiakirjaoid, syy_icd10id, ulksyy_icd10id,
dgdate, bdate, gender,
  incrementval, dmttype, rekisterinpitajaid, rekisteripitaja_organisaatioid
from (
select 1, dgid, pid, asiakirjaoid, syy_icd10id, ulksyy_icd10id,
dgdate, bdate, gender,
  row_number() over ( partition by pid order by dgdate ) as seq,
  '0' as dmttype, rekisterinpitajaid, rekisteripitaja_organisaatioid
from a_diagnosi dg
where ( syy_icd10id in ( select icd10id from dm_inclusion_icd10 )
or ulksyy_icd10id in ( select icd10id from dm_inclusion_icd10
) )
and enddate is null -- HUOM! skippaa diagnoosin lopetuskirjaukset
) where seq = 1;
logit('Number of patients identified by diagnosis: ' ||
TO_CHAR(SQL%ROWCOUNT), 'I');
update a_patient set dmttype = '1' where regexp_like( syy_icd10id,
'E10' ) or regexp_like( ulksyy_icd10id, 'E10' );
update a_patient set dmttype = '3' where dmttype='0' and regexp_like(
syy_icd10id, 'O24.4' ) or regexp_like( ulksyy_icd10id, 'O24.4' );
update a_patient set dmttype = '2' where dmttype='0' and regexp_like(
syy_icd10id, 'E11' ) or regexp_like( ulksyy_icd10id, 'E11' );
update a_patient set dmttype = '4' where dmttype='0' and regexp_like(
syy_icd10id, 'E1[234]' ) or regexp_like( ulksyy_icd10id, 'E1[234]'
);
end;

truncate table a_patient;
call update_patient();
```

Potilaan ryhmittely diabetestyyppeihin perustuu tässä siis potilaan ensimmäiseen kirjattuun diabetesdiagnoosiin. Diabetesdiagnoosit on kirjattu omaan aputauluunsa *dm\_inclusion\_icd10*:

```
create view dm_inclusion_icd10 as
select icd10id from d_icd10
where regexp_like( icd10id , 'E1[01234]|O24.4' );
```

Tietolähteeksi (sarake *src*) kirjataan arvo 1.

### Reseptikeskukseen kirjatut lääkemääräykset: ATC-koodi A10

Potilaiden taulua täydennetään lääkitystiedoilla:

```

merge into a_patient d using (
  select x.pid, prid, prdate
  from a_med x inner join a_prescription on cda_id = prid where
  medypeid = 1 ) s
on ( d.pid = s.pid )
when not matched then insert (pid, prid, src, dgdate) values
(s.pid, s.prid, 2, s.prdate )
when matched then update set d.src = bitor( d.src, 2 ), d.prid =
s.prid,
  d.dgdate = case when s.prdate < d.dgdate then s.prdate else
d.dgdate;

merge into a_patient d using (
  select pid, case
  when regexp_instr( atclist, 'A10B' ) > 0 and regexp_instr(
atclist, 'A10AB' ) > 0 and regexp_instr( atclist, 'A10A[CE]' ) > 0
then 22
  when regexp_instr( atclist, 'A10B' ) > 0 and regexp_instr(
atclist, 'A10A[CE]' ) > 0 then 21
  when regexp_instr( atclist, 'A10B' ) > 0 and regexp_instr(
atclist, 'A10AB' ) > 0 then 23
  when regexp_instr( atclist, 'A10B' ) > 0 and regexp_instr(
atclist, 'A10AD' ) > 0 then 24
  when regexp_instr( atclist, 'A10B' ) > 0 then 10
  when regexp_instr( atclist, 'A10A' ) > 0 and regexp_instr(
atclist, 'A10AB' ) > 0 and regexp_instr( atclist, 'A10A[CE]' ) > 0
then 33
  when regexp_instr( atclist, 'A10A' ) > 0 and regexp_instr(
atclist, 'A10A[CE]' ) > 0 then 31
  when regexp_instr( atclist, 'A10A' ) > 0 and regexp_instr(
atclist, 'A10AB' ) > 0 then 32
  when regexp_instr( atclist, 'A10A' ) > 0 and regexp_instr(
atclist, 'A10AD' ) > 0 then 34
  when regexp_instr( atclist, 'A10A' ) > 0 then 30 else 40 -- 40
should not appear
end as medypeid from (
  select pid, atclist from (
  select pid, listagg( substr(atc_code,1,5), '.') within group (
order by prdate ) as atclist
  from a_prescription inner join a_patient using (pid)
  where substr( atc_code, 1,3 ) = 'A10'
  group by pid )
) ) s
on ( d.pid = s.pid )
when matched then update set d.medypeid = s.medypeid;

```

Tämä lisäys perustuu tauluun *A\_Prescription*, joka on kuvattu myöhemmin kohdassa Lääkitys.

Huomattakoon, että on toinenkin määritelmä lääkinnällisestä alaryhmityksestä. Se perustuu potilaan viimeisimmän elinvuoden aikana **ostamiin** lääkkeisiin. Tämä sarake *medypeid2* määritellään vastaavalla tavalla, mutta rivit 23–25 korvataan seuraavasti:

```

select pid, listagg( substr(atc_code,1,5), '.') within group ( or-
der by dlvddate ) as atclist
  from a_delivery inner join a_patient using (pid)
  where coalesce( deathdate, to_date('31.12.2020')) - dlvddate <
366
  and substr( atc_code, 1,3 ) = 'A10'

```



Ks. taulu *A\_Delivery* jäljempää kohdasta Lääkitys.

### **bitor()-funktio**

Funktio *bitor()* on määritelty seuraavasti:

```
create function bitor( x integer, mask integer ) return integer as
begin
  return x + mask - bitand( x, mask );
end;
```

### **Henkilöt, joille etuusjärjestelmässä myönnetty lääkkeiden erityiskorvausoikeus 103 tai 215**

Erityiskorvausoikeudet päivitetään seuraavasti:

```
merge into a_patient d using (
  select pid, alpv from (
    select hexoraw(hetu) as pid, to_date( alpv, 'YYYY-MM-DD' ) as
    alpv,
    row_number() over (partition by hetu order by alpv) as seq
    from etuus_erytyiskorvausoikeus where sk1='103' )
  where seq = 1
) s
on ( d.pid = s.pid )
when not matched then insert ( pid, dgdate, src, reimbursement )
values ( s.pid, s.alpv, 4, 1 )
when matched then update set d.src = bitor( d.src, 4 ), d.reim-
bursement = 1,
  d.dgdate = case when s.alpv < d.dgdate then s.alpv else d.dgdate
end;

merge into a_patient d using (
  select pid, alpv from (
    select hexoraw(hetu) as pid, to_date( alpv, 'YYYY-MM-DD' ) as
    alpv,
    row_number() over (partition by hetu order by alpv) as seq
    from etuus_erytyiskorvausoikeus where sk1='215' )
  where seq = 1
) s
on ( d.pid = s.pid )
when not matched then insert ( pid, dgdate, src, reimbursement )
values ( s.pid, s.alpv, 4, 2 )
when matched then update set d.src = bitor( d.src, 4 ), d.reim-
bursement = d.reimbursement + 2,
  d.dgdate = case when s.alpv < d.dgdate then s.alpv else d.dgdate
end;
```

### **THL:n Hoitoilmoitusjärjestelmä: TerveysHilmo (ESH) ja AvoHilmo (PTH)**

Hilmo-päivitykset tapahtuvat seuraavasti:

```
-- Terveyshilmo (ESH)
merge into a_patient d using (
  select pid, tupva from (
    select hetu_aes as pid, tupva,
    row_number() over (partition by hetu_aes order by tupva) as seq
    from thl.tireka_hilmo_iso_icd10 inner join thl.tireka_hilmo_iso
    using (isoid)
    where regexp_like( koodi, 'E1[01234]|0244' )
  ) where seq = 1
) s
```

```

on ( d.pid = s.pid)
when not matched then insert ( pid, dgdate, src ) values ( s.pid,
s.tupva, 8)
when matched then update set d.src = bitor( d.src, 8 ),
d.dgdate = case when s.tupva < d.dgdate then s.tupva else
d.dgdate end;

--- Avohilmo ICD10 (PTH)
merge into a_patient d using (
select pid, kaynti_alkoi from (
select hetu_aes as pid, kaynti_alkoi,
row_number() over (partition by hetu_aes order by kaynti_alkoi)
as seq
from thl.tireka_avo_kayntisyy_icd10 inner join thl.ti-
reka_avo_palvelutapahtuma using (tapahtuma_tunnus)
where regexp_like( icd10, 'E1[01234]|O24.4')
) where seq = 1
) s
on ( d.pid = s.pid)
when not matched then insert ( pid, dgdate, src ) values ( s.pid,
s.kaynti_alkoi, 8)
when matched then update set d.src = bitor( d.src, 8 ),
d.dgdate = case when s.kaynti_alkoi < d.dgdate then s.kaynti_alkoi else d.dgdate end;
--- Avohilmo ICPC2 (PTH)
merge into a_patient d using (
select pid , kaynti_alkoi from (
select hetu_aes as pid, kaynti_alkoi,
row_number() over (partition by hetu_aes order by kaynti_alkoi)
as seq
from thl.tireka_avo_kayntisyy_icpc2 inner join thl.ti-
reka_avo_palvelutapahtuma using (tapahtuma_tunnus)
where regexp_like( icpc2, 'T89|T90|W85' )
) where seq = 1
) s
on ( d.pid = s.pid)
when not matched then insert ( pid, dgdate, src ) values ( s.pid,
s.kaynti_alkoi, 8)
when matched then update set d.src = bitor( d.src, 8 ),
d.dgdate = case when s.kaynti_alkoi < d.dgdate then s.kaynti_alkoi else d.dgdate end;

```

### Väestötietojärjestelmä

Lopuksi päivitetään Väestötietojärjestelmästä kaikille syntymäaika, sukupuoli, kuolinpäivä sekä viimeisin kotikunta.

```

merge into a_patient d using (
select hetu_aes, syntymapaiva, sukupuoli, kotikunta, kuolinpvm
from thl.tireka_henkilot
) s
on ( s.hetu_aes = d.pid )
when matched then update set d.bdate = syntymapaiva,
d.gender = case s.sukupuoli when '1' then 'mies' when '2' then
'nainen' else 'muu' end,
d.kuntaid = s.kotikunta, d.deathdate = s.kuolinpvm;

```

### Rekisteröintipäivä regdate

Sarake dgdate on edellä määritelty diagnoosin varhaisimmaksi toteamispäiväksi (*f\_diagnoosi-faktan toteamis\_pvm*). Jos kuitenkin jossain tietolähteessä esiintyy tätä aikaisempia päivämääriä, niin silloin sellaista on pidetty diagnoosipäivänä.

Elinaika-analyysien yhteydessä osoittautui, että hyvinvointialueilla vaikuttaa olevan erilaisia käytäntöjä varsinaisen diabetesdiagnoosipäivän toteamisessa. Sen vuoksi luotiin rinnakkainen "rekisteröintipäivämäärä" *regdate*, joka on yksinkertaisesti varhaisin päivämäärä, jolta missään käytössä olleista tietolähteistä henkilölle oli diabetekseen viittaavaa merkintää:

### Code Block 1 regdate

```
alter table a_patient add ( regdate date );
merge into a_patient d using (
  with dates as (
    select p.pid, trunc(asiakirjaluontiaika) as regdt
    from a_patient p inner join a_diagnoosi using (dgid)
    union
    select p.pid, prdate as regdt
    from a_patient p inner join a_prescription on prid = cda_id
    union
    select hextoraw(hetu) as pid, to_date( alp, 'YYYY-MM-DD' ) as
    regdt
    from etuus_erityiskorvausoikeus
  )
  select pid, min(regdt) as regdate
  from dates
  group by pid
) s
on ( d.pid = s.pid )
when matched then update set d.regdate = s.regdate;
```

## Lääkitys

Reseptikeskuksen aineisto (pilotissa taulu *Reseptiaineisto*) sisältää samassa tiedostossa niin lääkemääräykset, niiden korjaukset kuin toimituksetkin ja näiden korjaukset. Aineisto sisältää lisäksi myös sekä lääkemääräysten että niiden toimitusten mitätöintitapahtumat. Näitä tapahtumia ei pilotissa huomioitu.

Jotta erilaiset analyysit olisivat nopeampia ja sujuvampia tehdä, puretaan lääkemääräykset ja lääkkeiden toimitukset omiin tauluihinsa, *A\_Prescription* ja *A\_Delivery*, vastaavasti. Tauluja luotaessa pyrittiin korjaukset ottamaan huomioon, mutta yksinkertaisuuden vuoksi mitätöintitapahtumat jätettiin toistaiseksi huomiotta.

Reseptiaineiston primaariavain on *CDA\_ID*, joka on hajautustyyppinen surrogaattivain. Lääketoimitusdokumentit sisältävät viittauksen toimituksen pohjana olleen lääkemääräyksen *CDA\_ID*-tunnisteeseen.

### Lääkemääräykset *A\_Prescription*

```
create table a_prescription as
select hextoraw( cda_id ) as cda_id, hextoraw( patient_id ) as pid,
  atc_code, drug_name_c, organization_oid, dose_quantity_text,
  case when regexp_like( creation_date_fk, '\d{8}') then to_date(
  creation_date_fk, 'YYYYMMDD') else null end as prdate
from reseptiaineisto where doc_type_name_fi = 'Laakemaarays';

-- lääkemääräysten korjaukset
alter table a_prescription add constraint a_prescription_pk primary
key ( cda_id );

alter table a_prescription add ( new_id raw(16) );
```

```
merge into a_prescription d using (
  select hextoraw(cda_rplc_id) as cda_rplc_id, hextoraw( cda_id) as
  cda_id,
    hextoraw(patient_id) as pid, atc_code, drug_name_c, organiza-
    tion_oid,dose_quantity_text,
    case when regexp_like( creation_date_fk, '\d{8}') then to_date(
    creation_date_fk, 'YYYYMMDD') else null end as prdate
  from reseptiaineisto where doc_type_name_fi = 'Laakemaarayksen
  korjaus') s
on ( d.cda_id = s.cda_rplc_id )
when matched then update set d.new_id = s.cda_id, d.pid = s.pid,
d.atc_code = s.atc_code, d.drug_name_c = s.drug_name_c,
  d.organization_oid = s.organization_oid, d.dose_quantity_text =
  s.dose_quantity_text, d.prdate = s.prdate;

update a_prescription set cda_id = new_id where new_id is not null;
```

### Lääketoimitukset A\_Delivery

```
create table a_delivery as
select hextoraw( cda_id ) as cda_id, hextoraw( patient_id ) as pid,
  hextoraw( doc_group ) as prid, atc_code,
  case when regexp_like( creation_date_fk, '\d{8}') then to_date(
  creation_date_fk, 'YYYYMMDD') else null end as dlvdate,
  drug_name_c, organization_oid, dose_quantity_text, vnr_codel as
  vnr_code
from reseptiaineisto where doc_type_name_fi = 'Laakemaarayksen
toimitus';

-- lääketoimitusten korjaukset
alter table a_delivery add constraint a_delivery_pk primary key (
cda_id );

merge into a_delivery d using (
  select hextoraw(cda_rplc_id) as cda_rplc_id, hextoraw( cda_id) as
  cda_id, hextoraw( doc_group ) as prid,
    hextoraw(patient_id) as pid, atc_code, --drug_name_c, organiza-
    tion_oid,dose_quantity_text, vnr_codel,
    case when regexp_like( creation_date_fk, '\d{8}') then to_date(
    creation_date_fk, 'YYYYMMDD') else null end as dlvdate
  from reseptiaineisto where doc_type_name_fi = 'Laakemaarayksen
  toimituksen korjaus'
  order by creation_date_fk ) s
on ( d.cda_id = s.cda_rplc_id )
when matched then update d.pid = s.pid, d.prid = s.prid,
  d.drug_name_c = s.drug_name_c,d.organization_oid = s.organiza-
  tion_oid,
  d.dose_quantity_text = s.dose_quantity_text, d.vnr_code =
  s.vnr_codel
;
```

Hoitomyöntyvyyden arviointia helpottavan taulun *A\_Refill* rakentaminen on vielä kesken.

### Potilaiden lääkitys: taulu A\_Med

Taulu *A\_Med* kertoo, kenelle potilaista on määrätty verenpaine, lipidi tms. lääkitystä (ks. kohta Menetelmät, medtypeid=1-5). Se on luotu lääkemääräysten perusteella (taulu *A\_Prescription*):

## 6 Teknologiaaliite

```
create table a_med ( pid raw(16), medtypeid integer, cda_id
raw(16), dlvcount integer, recent_dlvcount integer,
firstdlv date, latestdlv date );
alter table a_med add constraint a_med primary key ( pid, medtypeid
);

insert into a_med ( pid, medtypeid, cda_id )
with data as (
select pid, cda_id, prdate,
case
when substr( atc_code, 1,3 ) = 'A10' then 1
when substr( atc_code, 1,3 ) between 'C02' and 'C09' then 2
when substr( atc_code, 1,3 ) = 'C10' then 5
when substr( atc_code, 1,4 ) = 'B01A' then 4
else 9 -- C01
end as medtypeid
from a_prescription
)
select pid, medtypeid, cda_id
from (
select pid, medtypeid, cda_id, row_number() over ( partition by
pid, medtypeid order by prdate ) as seq
from data
) where seq = 1;
-- C09 vielä lisäksi omaksi ryhmäkseen #3
insert into a_med ( pid, medtypeid, cda_id )
with data as (
select pid, cda_id, prdate, 3 as medtypeid
from a_prescription
where substr( atc_code, 1,3 ) = 'C09'
)
select pid, medtypeid, cda_id
from (
select pid, medtypeid, cda_id, row_number() over ( partition by
pid, medtypeid order by prdate ) as seq
from data
) where seq = 1;
```

Taulua päivitetään vielä ensimmäisen ja viimeisen lääkeitoimituksen tiedoilla viimeisten kahden vuoden ajalta (per 26.11.2020):

```
merge into a_med d using (
with data as (
select pid, dlvddate,
case
when substr( atc_code, 1,3 ) = 'A10' then 1
when substr( atc_code, 1,3 ) between 'C02' and 'C09' then 2
when substr( atc_code, 1,3 ) = 'C10' then 5
when substr( atc_code, 1,4 ) = 'B01A' then 4
else 9
end as medtypeid
from a_delivery
)
select pid, medtypeid,
count(*) as dlvcount, min(dlvddate) as firstdlv, max(dlvddate) as
latestdlv,
sum( case when dlvddate < current_date - 731 then 1 else 0 end )
as recent_dlvcount
from data
group by pid, medtypeid ) s
on ( d.pid = s.pid and d.medtypeid = s.medtypeid )
when matched then update set d.dlvcount=s.dlvcount, d.re-
cent_dlvcount = s.recent_dlvcount,
```

```

d.firstdlv=s.firstdlv, d.latestdlv=s.latestdlv;

-- C09:

merge into a_med d using (
with data as (
  select pid, cda_id, dlvdate, 3 as medtypeid
  from a_delivery
  where substr( atc_code, 1,3 ) = 'C09'
)
select pid, medtypeid,
  count(*) as dlvcnt, min(dlvdate) as firstdlv, max(dlvdate) as
latestdlv,
  sum( case when dlvdate < current_date - 731 then 1 else 0 end )
as recent_dlvcnt
from data
group by pid, medtypeid ) s
on ( d.pid = s.pid and d.medtypeid = s.medtypeid )
when matched then update set d.dlvcnt=s.dlvcnt, d.re-
cent_dlvcnt = s.recent_dlvcnt,
  d.firstdlv=s.firstdlv, d.latestdlv=s.latestdlv;

```

## Sairastavuus (comorbidity)

### a\_comorbidity

Tauluun *a\_comorbidity* kerättiin taulun *d\_disease* määrittelemät sairausedpisodit seuraavasti:

```

insert into a_comorbidity (pid, dgid, disdate, disid )
select distinct pid, dgid, dgdate, disid
from a_diagnoosi inner join dm_disease_icd10 on icd10id =
syy_icd10id;

insert into a_comorbidity (pid, procid, disdate, disid )
select distinct pid, procid, procddate, disid
from a_toimenpide inner join dm_disease_proc using (toimenpideluok-
kaid );

```

Tauluun on siis kirjattu henkilön **kaikki sairausedpisodit**, oli niiden taustana ICD10-indikaatio tai toimenpide.

Tästä käytännöstä horjuttiin myöhemmin lisäämällä vielä sairauskoodit 61–63, joihin onkin valittu potilaan **ensimmäinen** ao. sairaustapaus:

```

insert into a_comorbidity ( pid, disdate, disid, dgid, procid )
with diagnoses as (
  select pid, dgdate as disdate, disid, dgid, null as procid,
  row_number() over ( partition by pid order by dgdate ) as seq
  from a_diagnoosi join
  d_disease on (regexp_like( syy_icd10id, regexp_icd ) or re-
gexp_like( ulksyy_icd10id, regexp_icd ))
  where disid between 61 and 63
),
procedures as (
  select pid, procddate as disdate, disid, null as dgid, procid,
  row_number() over ( partition by pid order by procddate ) as seq
  from a_toimenpide inner join

```

```

    d_disease on regexp_like( toimenpideluokkaid, regexp_proc )
    where disid between 61 and 63
),
unioni as (
    select pid, disdate, disid, dgid, procid
    from diagnoses
    where seq = 1
    union
    select pid, disdate, disid, dgid, procid
    from procedures
    where seq = 1
)
select * from unioni x
    where not exists ( select 1 from unioni where pid = x.pid and
disid = x.disid and disdate < x.disdate )
;

```

### hilmo\_comorbidity

Jotta diagnoosien vertailu PTA:n ja Hilmon välillä olisi helppoa, luotiin Hilmoihiin kirjatusta ICD10-indikaatioista ja Terveyshilmon toimenpidekoodeista vastaava taulu seuraavasti:

```

create table hilmo_comorbidity as
select distinct 8 as src, disid, hetu_aes as pid, koodi as icd10id,
' ' as toimenpideluokkaid
from d_disease
    inner join thl.tireka_hilmo_iso_icd10 on regexp_like( koodi, rep-
lace( regexp_icd, '.', ''))
    inner join thl.tireka_hilmo_iso using (isoid)
;
insert into hilmo_comorbidity ( src, disid, pid, toimenpideluokkaid
)
    select distinct 8 as src, disid, hetu_aes as pid, tmpkoodi
    from thl.tireka_hilmo_iso_tmp inner join thl.tireka_hilmo_iso
using (isoid)
    inner join d_disease on regexp_like( tmpkoodi, regexp_proc )
;
insert into hilmo_comorbidity ( src, disid, pid, icd10id )
    select distinct 16 as src, disid, hetu_aes as pid, icd10
    from thl.tireka_avo_kayntisyys_icd10 inner join thl.ti-
reka_avo_palvelutapahtuma using (tapahtuma_tunnus)
    inner join d_disease on regexp_like( icd10, regexp_icd )
;

```

### r\_comorbidity

Nopean raportoinnin tueksi luotiin aggregaatiotaulu *r\_comorbidity*:

```

create table r_comorbidity as
select 8 as src, 'icd10' as ind, disid, title, koodi as code, icd10
as subtitle, extract( year from tupva ) as yr,
count(*) as n, count(distinct hetu_aes) as patients
from d_disease
    inner join thl.tireka_hilmo_iso_icd10 on regexp_like( koodi, re-
place( regexp_icd, '.', ''))
    inner join thl.tireka_hilmo_iso using (isoid)
    inner join a_patient on pid=hetu_aes and bitand( src, 1) = 1
left outer join v03_d_icd10 on koodi = replace( translate(
icd10id, '*+#. ', '____'), '_', '' )
group by disid, title, koodi, icd10, extract( year from tupva );
alter table r_comorbidity modify ( code varchar2(30) );

```

```

insert into r_comorbidity (src, ind, disid, title, code, subtitle,
yr, n, patients )
select 1 as src, 'icd10' as ind, disid, title, syy_icd10id as code,
      icd10 as subtitle, extract( year from disdate ) as yr,
      count(*) as n, count(distinct c.pid) as patients
from d_disease inner join a_comorbidity using (disid)
      inner join a_diagnoosi c using (dgid)
      left outer join v03_d_icd10 on syy_icd10id = icd10id
group by disid, title, syy_icd10id, icd10, extract( year from dis-
date );

insert into r_comorbidity (src, ind, disid, title, code, subtitle,
yr, n, patients )
select 1 as src, 'proc' as ind, disid, title, toimenpideluokkaid as
code,
      toimenpideluokka as subtitle, extract( year from disdate ) as yr,
      count(*) as n, count(distinct c.pid) as patients
from d_disease inner join a_comorbidity c using (disid)
      inner join a_toimenpide using (procid)
      left outer join v031_d_toimenpideluokka using (toimenpideluok-
kaid)
group by disid, title, toimenpideluokkaid, toimenpideluokka, ex-
tract( year from disdate );

insert into r_comorbidity (src, ind, disid, title, code, subtitle,
yr, n, patients )
select 8 as src, 'proc' as ind, disid, title, tmpkoodi as code,
toimenpideluokka as subtitle, extract( year from tupva ) as yr,
      count(*) as n, count(distinct hetu_aes) as patients
from d_disease
      inner join thl.tireka_hilmo_iso_tmp on regexp_like( tmpkoodi, re-
gexp_proc)
      inner join thl.tireka_hilmo_iso using (isoid)
      inner join a_patient on pid=hetu_aes and bitand( src, 1) = 1
      left outer join v031_d_toimenpideluokka on tmpkoodi = toimen-
pideluokkaid
group by disid, title, tmpkoodi, toimenpideluokka, extract( year
from tupva );

```

## Kuolleisuus

Kuolleisuuden analysointia varten luotiin näkymä *v\_survival* seuraavasti:

### Code Block 2 v\_survival

```

create or replace view v_survival as
select bdate, gender, dmtype, medtypeid2, hva, kuntaid,
      trunc(( to_date( '31.12.2020', 'DD.MM.YYYY' ) - bdate )/365.25)
as age,
      case when disid is null then 0 else 1 end as svt,
      case when bitand(src, 1 ) = 1 then 1 else 0 end as pta,
      case when bitand(src, 2 ) = 2 then 1 else 0 end as eresepti,
      case when bitand(src, 4 ) = 4 then 1 else 0 end as etuus,
      case when deathdate is null then 0 else 1 end as status,
      case when deathdate is not null then deathdate - regdate else
to_date('12.3.2021', 'DD.MM.YYYY' ) - regdate end as futime
from a_patient p inner join d_area using (kuntaid)
      left outer join a_comorbidity m on p.pid = m.pid and m.disid =
61
where regdate is not null and gender is not null and coalesce(

```



```
dmtype, '9' ) <> '3'
;
```

## Laboratoriotulokset

Laboratorioanalyysit, -nimikkeet ja -tulokset on 05-tason mallinnuksessa ripoteltu kaikki eri tauluihin. Tavoitteena on luoda yksi analyysikannan taulu A\_LAB, joka sisältää kunkin potilaan kaikkien laboratorioanalyysien arvot. Taulu muodostetaan seuraavasti:

### Code Block 3 v\_lab ja a\_lab

```
create or replace view v_lab as
with latest_version as (
  select * from (
    select hextoraw( f.potilashenkilotunnus ) as pid,
      asiakirjaoid, alkuperainenasiakirjaoid,
      asiakirjaluontiaika, rekisterinpitajapk, palvelutapahtumatun-
      nus, asiakirjaversio,
      row_number() over ( partition by alkuperainenasiakirjaoid order
by to_number(asiakirjaversio) desc ) as seq
    from f_laboratoriovastaus f
    inner join f_yhteinenheadertieto using (asiakirjaoid)
  )
  where seq=1
),
labanalysis as (
  select distinct asiakirjaoid, merkintaoid, entryoid,
    laboratoriotutkimusnimike, coalesce( tutkimuslyhenne,
cast('Pt-GFRe' as nvarchar2(10))) as tutkimuslyhenne
  from f_laboratoriotutkimusnimike n
  left outer join d_laboratoriotutkimusnimike d on d.laboratorio-
tutkimusnimikeid =
    case when n.paikallinentutkimusnimikeid = -1 then n.laborato-
riotutkimusnimikepk
    else n.paikallinentutkimusnimikeid end
)
select sys_guid() as labid, pid,
  asiakirjaoid, merkintaoid, entryoid,
  laboratoriotutkimusnimike, tutkimuslyhenne,
  case when regexp_like( tutkimustulosarvo, '\d+([\.,]\d+)?' )
    then to_number( replace( tutkimustulosarvo, '.', ',' ) ) else
  null end as value,
  tutkimustulosarvo, tutkimustulosyksikko, tutkimusaika,
  alkuperainenasiakirjaoid, asiakirjaversio, asiakirjaluontiaika,
  rekisterinpitajapk,
  palvelutapahtumatunnus, v.incrementval
from latest_version inner join f_laboratoriovastaus v using (asia-
kirjaoid)
  inner join labanalysis using (asiakirjaoid, merkintaoid, entryoid)
)
  left outer join f_laboratoriotutkimustulos using (asiakirjaoid,
merkintaoid, entryoid )
;
create table a_lab as select * from v_lab;
```

Taululle luodaan surrogaattivain *labid*, mutta luonnollinen avain koostuisi nelikosta (*asia-  
kirjaoid, merkintaoid, entryoid, tutkimuslyhenne*). Tämän asettaminen ei kuitenkaan onnistu,

koska on olemassa kourallinen havaintoja (29.8.2021 yhteensä 26 duplikaattia), joissa samalle tutkimuslyhenteelle löytyy kaksi arvoa. Yleensä näissä on kyse siitä, että arvo on ilmaistu kahdella eri mittausyksiköllä. Mutta aineistosta löytyi myös yksi tapaus, jossa yksikkökin oli sama, vain arvo vaihteli.

Aineistossa on arvoja, jossa kenttään *paikallinentutkimusnimikeid* oli kirjattu arvo 10457. Näille on tutkimuslyhenteeksi asetettu eo. skriptissä 'Pt-GFRE'.

### Tekstivastaukset

Edellä kuvattu näkymän *V\_Lab* määrittely poimii analyysien numeeriset tulosarvot faktataulusta *F\_LaboratoriotutkimusTulos*. Sen sarake *Tutkimustulosarvo* sisältää kuitenkin vain sanan 'Puuttuu', jos arvoa ei ole ilmoitettu numeerisena. Itse asiassa näissäkin tapauksissa faktataulun *F\_LaboratorioVastaus* sarake *Tutkimustulosteksti* voi sisältää tutkimuksen tuloksen tekstimuodossa. Melko yleinen käytäntö on, että tuloksen perään on kirjattu tähti kuvaamaan raja-arvojen ylittävää mittaustulosta. On myös mahdollista, että tulos on kirjattu sanallisena tyyliin "Lopullinen tulos 40 mmol/mol". Tämän vuoksi analyysitauluun *A\_Lab* lisättiin vielä kenttä *Tutkimustulosteksti*, josta puolestaan säännöllisellä lausekkeella poimittiin numeeriset tulosarvot sarakkeeseen *Value* seuraavasti:

```
-- Uusi sarake "tutkimustulosteksti"
alter table a_lab add ( tutkimustulosteksti nvarchar2(1020));
merge into a_lab d using (
  select labid, f.tutkimustulosteksti
  from a_lab inner join v031_f_laboratoriovastaus f using (asiakir-
jaoid, merkintaoid, entryoid )
  where f.tutkimustulosteksti is not null
) s
on ( d.labid = s.labid )
when matched then update set d.tutkimustulosteksti = s.tutkimustu-
losteksti;

-- Sarakkeen "value" päivitys säännöllisellä lausekkeella
merge into a_lab d using (
  select labid,
  to_number(replace(
    regexp_substr( regexp_substr( tutkimustulosteksti,
'\d+([\.,]\d+)?\s*(\*|mmol|ml|mg|umol|umol|ug)', 1, 1, 'i'),
'\d+([\.,]\d+)?'),
  '.', ',' )) as value
  from a_lab
  where value is null
  and regexp_like( tutkimustulosteksti,
'\d+\s*(\*|mmol|ml|mg|umol|umol|ug)', 'i')
) s
on (d.labid = s.labid )
when matched then update set d.value = s.value;

-- HbA1c-prosenttiarvojen päivitys:
merge into a_lab d using (
  select labid,
  to_number(replace(
    regexp_substr( regexp_substr( tutkimustulosteksti,
'\d+([\.,]\d+)?\s*%'), '\d+([\.,]\d+)?'),
  '.', ',' )) as value
  from a_lab
  where tutkimuslyhenne = 'B -HbA1c' and value is null
  and regexp_like( tutkimustulosteksti, '\d+\s*%')
) s
on (d.labid = s.labid )
when matched then update set d.value = 10.93*s.value-23.5, d.tutki-
mustulosyksikko = 'mmol/mol*';
```

Skriptin lopussa on myös mahdollisesti prosenttisarvoina ilmoitettujen HbA1c-tulosten lisäys ja konvertointi mitta-asteikolle mmol/mol.

### Monistuneet rivit

Osoittautui, että tauluun *a\_lab* kertyy edellä kuvatulla algoritmilla vielä useita rivejä, joissa on mitä ilmeisimmin kyse samasta analyysistä: samalle henkilölle on samaan tutkimusaikaan (joka on aikaleima) kirjattu sama analyysi – joko samalla tuloksella tai eri vastausarvolla – useita kertoja, pahimmillaan 199 kertaa. Näiden *alkuperainenasiakirjaoid*-tunnisteet (ns. *setld*) kuitenkin eroavat toisistaan. Tällaisia tapauksia oli yli 60000 kappaletta (0,7 % kaikista tapauksista). Näissä tapauksissa saattoi jopa asiakirjan luontiaikakin olla sekunnilleen sama. Selvää syytä asialle ei ole keksitty.

Ylimääräiset rivit poistettiin seuraavalla skriptillä:

```
delete from a_lab where labid in (
  select labid from (
    select labid, row_number() over (partition by pid, tutkimus-
lyhenne, tutkimusaika order by asiakirjaluontiaika desc) as seq
    from a_lab
  ) where seq > 1
);
```

Kaikki monistuneet rivit tallennettiin myöhempiä analyysyjä varten tauluun *tmp\_lab\_dup* ennen rivien poistamista.

### Pt-GFReEPI

[Glomerulusten suodatusnopeus \(GFR\) arvioituna CKD-EPI-kaavalla](#) laskettiin tauluun *a\_lab* seuraavasti:

```
insert into a_lab (labid, pid, asiakirjaoid, merkintaoid, entryoid,
tutkimuslyhenne, tutkimusaika, rekisterinpitajapk, incrementval,
value )
with data as (
  select pid, gender, trunc((trunc(tutkimusaika) - bdate)/365.25)
as age, value,
  x.asiakirjaoid, x.merkintaoid, x.entryoid, x.tutkimusaika,
x.rekisterinpitajapk
  from a_patient inner join a_lab x using (pid)
  where tutkimuslyhenne = 'P -Krea'
)
select sys_guid(), pid, asiakirjaoid, merkintaoid, entryoid, 'Pt-
GFReEPI', tutkimusaika, rekisterinpitajapk, 0,
  case when gender='nainen' and value <= 62 then 144*power(va-
lue/61.9, -0.329)*power(0.993,age)
  when gender='nainen' and value > 62 then 144*power(va-
lue/61.9, -1.209)*power(0.993,age)
  when gender='mies' and value <= 80 then 141*power(va-
lue/79.6, -0.411)*power(0.993,age)
  when gender='mies' and value > 80 then 141*power(value/79.6,
-1.209)*power(0.993,age)
  else null
  end
  from data;
```

## Fysiologiset mittaukset

Fysiologisten mittausten taulu *A\_FysiologinenMittaus* muodostetaan seuraavasti:

```

create or replace view v_fysiologinenmittaus as
with latest_version as (
  select * from (
    select hextoraw( x.potilashenkilotunnus ) as pid,
      asiakirjaoid,
      row_number() over ( partition by alkuperainenasiakirjaoid order
by to_number(asiakirjaversio) desc ) as seq
    from f_fysiologinenmittaus x inner join f_yhteinenheadertieto
using (asiakirjaoid)
  )
  where seq=1
)
select sys_guid() as fid, pid,
  fysiologinenmittausid, fysiologinenmittaus, mittaasteikko,
  mittausaika, mittaustulosnumeerinen, mittaustulosyksikko, mit-
taustulostekstina,
  asiakirjaoid, merkintaoid, entryoid
from f_fysiologinenmittaus x inner join latest_version using (asia-
kirjaoid)
  left outer join d_fysiologinenmittaus d on x.fysiologinenmit-
tauspk = d.fysiologinenmittausid
;

create table a_fysiologinenmittaus as select * from v_fysiologinen-
mittaus;
alter table a_fysiologinenmittaus add constraint a_fysiologinenmit-
taus_pk primary key (fid);

```

## Tukiaineistot

### Julkisten rekisterinpitäjien PTA-laboratorioanalyysit

Aggregaatti "aggroidut\_labrat" muodostettiin periaatteessa seuraavankaltaisella SQL-lauseella 05-tason Spark-tietokannasta. Tässä esitettävä kysely vastaa 05-mallinnuksen Oracle-syntaksia:

```
create table aggroidut_labrat as
select
  h.rekisterinpitajapk as rekisterinpitajaid,
  tietojarjestelmanimi,
  coalesce( nullif( laboratoriotutkimusnimikepk, cast('-1' as nvarchar2(2)) ), paikallinentutkimusnimikeid ) as nimikeid,
  paikallinentutkimusnimike as nimike,
  count(*) as doccount, count(distinct alkuperainenasiakirjaoid) as
setcount,
  min(trunc(asiakirjaluontiaika)) as mindate, max(trunc(asiakirjaluontiaika)) as maxdate
from f_laboratoriotutkimusnimike inner join f_yhteinenheadertieto h
using (asiakirjaoid)
  inner join d_rekisterinpitaja on h.rekisterinpitajapk = rekisterinpitajaid
-- where asiakirjaluontiaika >= '1.1.2019'
group by coalesce( nullif( laboratoriotutkimusnimikepk, cast('-1' as nvarchar2(2)) ), paikallinentutkimusnimikeid ),
  paikallinentutkimusnimike, h.rekisterinpitajapk, tietojarjestelmanimi
;
```

Myöhemmin muodostettiin myös toinen, spesifimpi aggregaatio "aggroidut\_labrat\_2019", joka keskittyy vuonna 2019 kirjattuihin laboratoriotuloksiin (vrt. eo. koodin rivi 11).