



HELSINGIN YLIOPISTO

**Työkyvyttömyyden pitkittyminen julkisella sektorilla:  
Työkyvyttömyyseläkkeiden lääketieteellisen syyn  
vaihtuminen tuki- ja liikuntaelinsairauksista  
psykiatrisiksi sairauksiksi**

Helsingin yliopisto  
Valtioteellinen tiedekunta  
Sosiaalitieteiden maisteriohjelma  
Yhteiskuntapolitiikka  
Maisterintutkielma

Laatija:  
Jesse Vinnari

Ohjaaja:  
Dos. Janne Salonen

1.11.2024  
Helsinki

**Tiedekunta:** Valtiotieteellinen tiedekunta, Helsingin yliopisto

**Koulutusohjelma:** Sosiaalitieteiden maisteriohjelma

**Opintosuunta:** Yhteiskuntapolitiikka

**Tekijä:** Jesse Vinnari

**Työn nimi:** Työkyvyttömyyden pitkittyminen julkisella sektorilla:

Työkyvyttömyyseläkkeiden lääketieteellisen syyn vaihtuminen tuki- ja liikuntaelinsairauksista psykiatrisiksi sairauksiksi

**Työn laji:** Maisterintutkielma

**Kuukausi ja vuosi:** Marraskuu 2024

**Sivumäärä:** 80 + lähteet ja liitteet (yht. 94 sivua)

**Avainsanat:** Työkyvyttömyyseläke, sekvenssianalyysi, kvantitatiivinen tutkimus, julkinen sektori, työkyvyttömyys, diagnoosin vaihtuminen, tuki- ja liikuntaelinsairaudet, psykiatrian sairaudet, hakemuspolut

**Ohjaaja tai ohjaajat:** Dos. Janne Salonen

**Säilytyspaikka:** Helsingin yliopiston kirjasto – Helda / E-thesis (opinnäytteet)

**Muita tietoja:**

**Tiivistelmä:**

Työkyvyttömyys on haastava käsite ja tutkimuksellisesti vaativa kohde. Kun työkyvyttömyyttä tarkastellaan työkyvyttömyyseläkkeiden näkökulmasta, tulee ilmiöistä erityisen monisyinen. Kuitenkaan työkyvyttömyyden seurauksista yksilöille tai yhteisöille ei ole epäselvyyttä: työkyvyttömyys on yksilölle henkilökohtainen tragedia, yhteisölle menetetty resurssi, ja molemmille suuria kustannuksia aiheuttava epätoivottu tapahtuma. Suomessa vuonna 2023 työeläkejärjestelmästä työkyvyttömyyseläkkeelle siirtyi 17 700 henkilöä. Tämän joukon tyypillisimmät työkyvyttömyyden syyt olivat tuki- ja liikuntaelinsairaudet sekä psykiatrian sairaudet. Tässä maisterintutkielmassa päätarkoituksena oli tuottaa uutta tietoa työkyvyttömyyseläketutkimukseen Suomessa julkisella sektorilla. Tarkastelun

kohteena olivat tuki- ja liikuntaelinsairauksista psykiatrian sairauksiksi ensisijaiselta lääketieteelliseltä diagnosoiltaan vaihtuneet työkyvyttömyyseläkeratkaisut.

Tutkielman aineistona käytettiin Kevan rekisteriaineistoa työkyvyttömyyseläkeratkaisuksista vuosina 2013–2023. Aineistossa oli yhteensä 61 041 henkilöä, joista monilla oli useita ratkaisuja. Tutkimusaineistosta poimittiin tuki- ja liikuntaelinsairautena alkaneita ratkaisuja, ja tutkittiin sekvenssianalyysillä hakemuspolkuja tarkoituksena selvittää missä määrin ensisijainen lääketieteellinen syy vaihtui psykiatrian sairauksiksi, sekä tarkastella miten muutos näkyi iän ja sukupuolen suhteen. Sekvenssianalyysi on luonteeltaan eksploratiivinen, ja menetelmän tarkoituksena oli pyrkiä kartoittamaan ja ymmärtämään ilmiötä avoimella ja joustavalla tavalla. Tavoitteena oli löytää uusia näkökulmia, ideoita ja hypoteeseja myöhempää, tarkempaa tutkimusta varten.

Tulokset osoittivat, että julkisella sektorilla on olemassa selkeä väestöryhmä, jolla tuki- ja liikuntaelinsairautena alkaneet työkyvyttömyyseläkeratkaisut vaihtuivat psykiatrian sairauksiksi, ja vaihtuminen tapahtui keskimäärin kolmen vuoden sisällä ensimmäisestä työkyvyttömyyseläkeratkaisusta. Tämä vaihtuminen oli tyypillisempää nuoremmilla ikäryhmillä ja naisilla. Tulosten perusteella voitiin myös todeta, että sekvenssianalyysi on käyttökelpoinen menetelmä tämänkaltaisessa yhteiskuntapoliittisessa tutkimuksessa, ja työkyvyttömyyseläkkeiden hakemuspolkuja tutkimalla voidaan saada uutta tietoa ja jatkotutkimuksen aiheita.

Tutkielman tulokset ovat linjassa aikaisemman tutkimuksen kanssa. Vastaavaa menetelmää ja tutkimusasetelmaa ei parhaan tietoni mukaan ole käytetty aiemmassa työkyvyttömyystutkimuksessa, joten tämä tutkielma on tuottanut uutta tietoa tutkimuskentälle. Suomalainen työkyvyttömyyseläkejärjestelmä tuo turvaa, mutta siinä on myös kehitettävää. Tulevaisuudessa työkyvyttömyyden ehkäisy on erityisen tärkeässä roolissa niin yksilön kuin yhteisönkin näkökulmasta, ja tutkimuksellista katsetta on syytä tarkentaa monipuolisesti eri työkyvyttömyyden alueille.

# Sisällysluettelo

<b>1</b>	<b>Johdanto</b>	<b>1</b>
<b>2</b>	<b>Aiempi tutkimus</b>	<b>6</b>
2.1	Kansainvälinen tutkimus	7
2.2	Suomalainen tutkimus	10
<b>3</b>	<b>Tutkimuskysymys, käsitteet ja teoreettinen viitekehys</b>	<b>15</b>
3.1	Tutkimuskysymys	15
3.2	Käsitteet	16
3.3	<b>Eläke- ja työkyvyttömyyseläketutkimuksen teoreettinen viitekehys</b>	<b>19</b>
3.3.1	Teoreettiseen viitekehykseen liittyvä taustatutkimus	19
3.3.2	Eläkkeelle siirtymisikä ja terveys	20
3.3.3	Eläkkeelle jäämisen negatiiviset seuraukset	22
3.3.4	Työolojen, kuormituksen, sairauksien ja eläkkeen yhteys työkyvyttömyyteen	23
3.4	<b>Sekvenssianalyysin teoreettinen viitekehys</b>	<b>26</b>
<b>4</b>	<b>Aineisto ja menetelmä</b>	<b>29</b>
4.1	<b>Aineisto</b>	<b>29</b>
4.2	<b>Menetelmä</b>	<b>31</b>
4.2.1	Sekvenssianalyysi	31
4.2.2	Sekvenssin Alphabet	33
4.2.3	Puuttuvat arvot	34
4.2.4	Optimal Matching -menetelmä	35
4.2.5	Substituutiomatriisi ja TRATE	36
4.2.6	Visualisointi	37
4.2.7	Objektit ja ryhmittely (klusterointi)	38
4.2.8	Sekvenssianalyysin ja ryhmittelyn (klusterointi) laadunarviointi ja vertailu	40
4.3	<b>Tutkimusetiikka</b>	<b>42</b>
<b>5</b>	<b>Tulokset</b>	<b>43</b>
5.1	<b>Aineiston yleisesittely</b>	<b>43</b>
5.2	<b>Sekvenssianalyysi hakemuspoluista</b>	<b>54</b>
5.3	<b>Hakemuspolkujen ryhmittely (klusterointi)</b>	<b>58</b>
5.3.1	Ryhmittelyn päätulokset	59
5.3.2	Ikä ja sukupuoli ryhmissä	64

<b>5.4</b>	<b>Laadunarviointi ja vertailu</b>	<b>67</b>
<b>6</b>	<b>Pohdinta</b>	<b>70</b>
<b>6.1</b>	<b>Keskeiset tulokset</b>	<b>71</b>
6.1.1	Sukupuolten erot	72
6.1.2	Ikäryhmien erot	73
<b>6.2</b>	<b>Tutkimuksen luotettavuus ja kriittinen tarkastelu</b>	<b>74</b>
<b>7</b>	<b>Johtopäätökset</b>	<b>78</b>
	<b>Lähteet</b>	<b>81</b>
	<b>Liitteet</b>	<b>89</b>

## 1 Johdanto

Työkyvyttömyys on haastava käsite ja tutkimuksellisesti vaativa kohde. Kun työkyvyttömyyttä tarkastellaan työkyvyttömyyseläkkeiden ja sen kaltaisten etuuksien näkökulmasta, tulee ilmiöistä erityisen monisyinen. Kuitenkaan työkyvyttömyyden seurauksista yksilöille tai yhteisöille ei ole epäselvyyttä: työkyvyttömyys on yksilölle henkilökohtainen tragedia, yhteisölle menetetty resurssi, ja molemmille yleensä suuria haittoja ja kustannuksia aiheuttava epätoivottu tapahtuma. Tässä maisterintutkielmassa päätarkoituksena on tuottaa uutta tietoa työkyvyttömyyseläketutkimukseen Suomessa julkisella sektorilla. Tarkastelun kohteena on Kevan työkyvyttömyyseläkeratkaisut rekisteriaineistosta kymmenen vuoden ajalta. Menetelmänä on sekvenssianalyysilla ja klusterianalyysilla tuotetut hakemuspolut, joista huomion kohteena ovat tuki- ja liikuntaelinsairauksista psykiatrian sairauksiksi ensisijaiselta lääketieteelliseltä diagnoosiltaan vaihtuneet työkyvyttömyyseläkeratkaisut.

Kuten muualla Euroopassa, niin myös Suomessa väestö vanhenee matalan syntyvyyden ja pidentyvän eliniän myötä. On arvioitu, että vuonna 2050 lähes 30 prosenttia Euroopan väestöstä on yli 65-vuotiaita ja vanhushuoltosuhte tulee kasvamaan. Euroopan maissa lähes puolella 55–64-vuotiaista on jokin pitkäaikainen sairaus tai terveysongelma, ja huono terveys on yksi keskeisimmistä syistä varhaiseen poistumiseen työmarkkinoilta. Samalla monet sairaudet ovat muuttuneet aiempaa vähemmän rajoittaviksi. Tähän vaikuttavat muun muassa fyysisen työn väheneminen teknologisen kehityksen vuoksi ja työpaikkojen siirtyminen palvelualalle. Toisaalta henkinen kuormitus ja kognitiivisen työn vaatimukset ovat lisääntyneet nostoen uudenlaisia syitä työkyvyttömyyden taustalle. Lääketieteellisten hoitojen kehitys ja työllisyyttä edistävän työkyvyttömyyspolitiikan tehostaminen voivat tuottaa lisää mahdollisuuksia jatkaa työssä työkyvyn alenemisesta huolimatta. (Laaksonen ym. 2022.)

Eläketurvakeskuksen tilastojen mukaan Suomessa vuonna 2023 työeläkejärjestelmästä työkyvyttömyyseläkkeelle siirtyi 17 700 henkilöä, joista 5 700 henkilöllä ensisijainen työkyvyttömyyden syy oli mielenterveyden ja käyttäytymisen häiriö. Kaikista vuonna 2023 työkyvyttömyyseläkkeelle siirtyneistä 8 400 oli miehiä ja 9 300 naisia. Miehillä tyypillisin syy oli tuki- ja liikuntaelinsairaudet (32 %) ja

toiseksi tyypillisin syy oli mielenterveyden ja käyttäytymisen häiriöt (26 %), kun vastaavasti naisilla tyypillisin syy oli mielenterveyden ja käyttäytymisen häiriöt (37 %) ja toiseksi yleisimpänä syynä oli tuki- ja liikuntaelinsairaudet (32 %). Lisäksi alle 55-vuotiaiden, ja erityisesti alle 35-vuotiaiden, työkyvyttömyyshaasteet ovat lisääntyneet runsaasti 2000-luvulla. (ETK Työkyvyttömyyseläkkeelle 2024).

Vuodesta 2008 alkaen 16–34-vuotiaiden työkyvyttömyyseläkkeiden määrä on kasvanut joka vuosi 2020 vuoteen saakka, ja vaikka tämän jälkeen määrä on hieman alkanut laskea, niin alle 35-vuotiaiden työkyvyttömyyseläkkeillä olevien määrä oli vuoden 2022 lopussa yli 25 000 henkilöä (KELA Työkyvyttömyyseläke 2024).

Vuonna 2022 sairauspäivärahaa sai tuki- ja liikuntaelinsairauden vuoksi 79 400 ja psykiatriansairauden vuoksi 98 000 henkilöä. Jos henkilöllä oli molemmat diagnoosit, niin sairauspoissaoloja oli tilastollisesti merkitsevästi enemmän, ja myös terveydenhuollon kustannuksia aiheutui selvästi enemmän.

Työkyvyttömyyseläkkeellä olevat menettävät työansioita noin 2,5 miljardia euroa vuodessa ja menetetyn työpanoksen arvo on vähintään 3 miljardia euroa. Pelkästään mielenterveyden kustannukset Suomelle ovat 11 miljardia euroa vuodessa. Tuki- ja liikuntaelinsairauksista vastaavaa arviota ei ole tehty. (Ryynänen & Metsä 2024.)

Työkyvyttömyyden haasteet korostuvat erityisesti julkisella sektorilla, kun työvoima ikääntyy ja tarpeet työkyvyn ylläpitämiselle muuttuvat, eikä näistä haasteista aiheutuvat kustannukset ole luultavasti lähiaikoina kääntymässä laskuun. Julkisella sektorilla työkyvyttömyyden on arvioitu aiheuttavan vuosittain yli 200 000 henkilön työpanoksen menetyksen, ja työkyvyttömyyden alentaminen nykytasolta 20 prosenttia voisi nostaa työllisyysasteen jopa 76 prosenttiin (Telma Kuinka kalliiksi työkyvyttömyys tulee 2023). Sosiaali- ja terveysministeriö arvioi menetetyn työpanoksen hinnaksi jopa kahdeksan miljardia euroa (Rissanen & Kaseva 2014). Tekemättömän työn vuosikatsauksessa 2016 arvioitiin työkyvyttömyyden kokonaiskustannusten aiheuttavan arviolta noin neljän miljardin euron kustannuksen yrityksille ja yhden miljardin kustannuksen kunnille (Tekemättömän työn vuosikatsaus 2016). Vuonna 2023 julkaistun artikkelin mukaan menetetyn työpanoksen kustannukset olivat vähintään kolme miljardia euroa vuonna 2021 (Lappo & Risku 2023). Työkyvyttömyyteen liittyvissä menetyksissä puhutaan merkittävistä resursseista.

Mielenterveyden häiriöt ovat erittäin yleisiä ja inhimillistä kärsimystä aiheuttavia sairauksia, ja ne esiintyvät usein yhtä aikaa muiden sairauksien kanssa.

Mielenterveyden häiriöiden kustannuksia voidaan mitata terveydenhuollon kustannuksilla, tuottavuuden laskulla, sairauspoissaoloilla ja inhimillisellä kärsimyksellä. (Goodell ym. 2011.) Terveyden ja mielenterveyden vahvan yhteyden vuoksi mielenterveyden vaikutukset on otettava huomioon, kun tarkastellaan fyysistä terveyttä ja terveyden sosiaalisia määrittäjiä tarkemmin (Compton & Shim 2015, 424). Fyysisten työolojen muutosten on havaittu olevan erityisesti keski-ikäisillä palkansaajilla yhteydessä työkyvyn kehitykseen (Laaksonen 2016). Kuitenkaan työkyvyn alenemat eivät välttämättä aina tule näkyviin, sillä työntekijät saattavat sinitellä työssään ongelmien kanssa. Laaksonen ym. (2022) osoittavat tutkimuksessaan, että Suomessa työntekijöiden työssäolovuodet ovat viime vuosina lisääntyneet siksi, että terveysongelmat ovat vähentyneet ja työkyky parantunut, mutta myös siksi, että terveysongelmien ja huonon työkyvyn kanssa työskentely on lisääntynyt. Töissä yritetään selviytyä.

Työkyvyn ja työkyvyttömyyden rajapinnoilla on paljon ihmisiä, joiden tilannetta parantamalla voisi olla mahdollista kohentaa sekä yksilöllistä että yhteiskunnallista tuottavuutta ja lisätä hyvinvointia. Tämä tarkoittaa esimerkiksi määräaikaista ja osa-aikaista työkyvyttömyyksiä, joissa työkyvyn alenema ei ole pysyvää.

Työkyvyttömyyseläkkeiden säätely on kuitenkin monimutkaista, ja voi helposti syntyä tilanteita, joissa työskentely työkyvyttömyyseläkkeellä muodostuu pelkkää työntekoa houkuttelevammaksi vaihtoehdoksi. Usein työkyvyttömyyseläke ei myöskään ole ainoa henkilön saama sosiaaliturvamuoto, jolloin yhden sosiaaliturvan osan muokkaaminen voi aiheuttaa mittavia välillisiä muutoksia kokonaisuudessa, ja pahimmillaan yksittäinen parannus voi olla kokonaisuudessa heikennys. Toisinaan myös työkyvyttömyyseläkkeisiin liittyvät ehdot tai rajoitukset voivat olla epäedullisia työntekijälle, sillä ne eivät jousta yhteiskunnassa tapahtuvien muutosten, kuten inflaation, mukana (STM 2017). Työkyvyttömyyteen liittyvien ilmiöiden ymmärtämiseksi on olennaista selvittää, miksi ja kuinka laajasti työkyvyttömyyttä esiintyy. Ilmiöiden tarkempi analysointi on mahdollista silloin, kun tunnistetaan työkyvyttömyyden taustalla olevat syyt, erityisesti tilanteissa, joissa työkyvyttömyys pitkittyy tai sen syyt muuttuvat.

Koska työkyvyttömyys on ilmiönä monisyinen, rajataan tässä maisterintutkielmassa taustateoriaa tutkimuskysymykselle ja aineistolle sopivalla tavalla, eikä pyritä tyhjentävästi vastaamaan kaikkiin työkyvyttömyystutkimuksen haasteisiin. Garrouste ja Perdrix (2022) taustoittavat laajan meta-analyysinsa tuloksena kolme tekijää, jotka tuovat työkyvyttömyyseläketutkimuksenkin haasteita näkyviksi. Ensinnäkin, koska on hyvin vaikea verrata eläkkeelle jäämisen ja fyysisen terveyden yhteyttä yleisesti, tutkimustuloksia on tehtävä paikallisesti ja harmonisoitava, jotta on mahdollista ymmärtää eroja. Toiseksi tutkimukset, jotka eivät löydä tilastollisesti merkitsevää vaikutusta, eivät aina suoraan voi osoittaa onko kyseessä todella vaikutuksen puuttuminen vai kyvyttömyyttä havaita vaikutus. Siksi on tarpeellista tehdä tutkimusta myös työkyvyttömyyteen ilmiönä sopivista vaikeasti havaittavista vaikutuksista, ja sitä kautta ymmärtää, missä tilastollisen voiman puute todella on. Kolmanneksi vain harvat viimeaikaiset eläketutkimukset ovat osoittaneet vaikutusta monipuolisesti, ja yksityiskohtaisempaa tutkimusta juuri reuna-alueilla tarvitaan. Näiden kolmen tuloksen pohjalta ja mainittuihin tekijöihin nojaten tässä maisterintutkielmassa pyritään tuomaan eksploraatiivisilla menetelmillä näkökulmia alan tutkimukseen.

Tässä työssä sivutaan myös lääketiedettä käsiteltäessä työkyvyttömyyden ensisijaisia syitä työkyvyttömyyseläkeratkaisuissa, mutta työn näkökulma on yhteiskuntapoliittinen. Taustalla on pyrkimys selvittää, miten voidaan parantaa ihmisten työkykyä ja julkisen sektorin organisaatioiden ymmärrystä työkyvyttömyydestä, jotta hyvinvointia olisi mahdollista lisätä niin yksilö- kuin yhteiskuntatasollakin. Prosessina Suomessa julkisella sektorilla työkyvyttömyyttä hakeva työntekijä on kohdannut sekä lääkärin että asiantuntijalääkärin, ja Kevan ratkaisuasiantuntija on näiden antamien lausuntojen perusteella tehnyt ratkaisun oikeudesta työkyvyttömyyseläkkeeseen.

Kevan lakisääteinen tehtävä eläkkeiden lisäksi on hoitaa ”julkisten alojen eläkelain mukaisen kunta-alan, hyvinvointialueiden, hyvinvointiyhtymien, valtion, evankelis-luterilaisen kirkon, Kansaneläkelaitoksen ja Suomen Pankin henkilöstön eläketurvan toimeenpanoa ja henkilöstön työkyvyttömyysriskin vähentämiseen liittyvää toimintaa sekä muita mainitussa laissa säädettyjä tehtäviä” (Laki Kevasta 66/2016). Keva toteuttaa tehtävää tarjoamalla työnantajille työelämäpalveluita ja tutkittua tietoa työkykyjohtamisen tueksi. Työkyvyn pysyvään menettämiseen liittyvä työkyvyn

heikkeneminen on useimmilla henkilöillä pitkäaikainen prosessi. Käytännössä hakija on aloittanut polkunsä hoitavan lääkäriin ja työterveyshuollon kanssa, ja työkyvyttömyyttä edeltää pitkä sairauspoissaolojakso. Jos työkyky on alentunut riittävästi, niin henkilö asioi eläkelaitoksen vakuutuslääkäriin kanssa. Tämä maisterintutkielma lukeutuu osaksi Kevan työkyvyttömyysriskin vähentämiseen liittyvää tehtävää ja tutkimustyötä.

Tässä työssä noudatetaan parhaita tunnettuja käytäntöjä aiheen ja tieteenalan erityispiirteiden osalta. Tutkimuskentän haasteet on tunnistettu koskien suomalaista ja kansainvälistä tutkimusta, ja niiden soveltamista suomalaiseen rekisteriaineistoon. Tutkimusaineistona käytetään Kevan rekisteriaineistosta julkisen sektorin työkyvyttömyyseläkeratkaisuja vuosilta 2013–2023. Tilastollisena tutkimusmenetelmänä hyödynnetään eksploratiivista sekvenssimenetelmää, jonka avulla voidaan tarkastella rekisteriaineistosta työkyvyttömyyseläkeratkaisujen ensisijaisen lääketieteellisen syyn vaihtumista tuki- ja liikuntaelinsairaudesta psykiatriseksi sairaudeksi työkyvyttömyyden jatkuessa. Lisäksi tarkastellaan missä määrin tätä vaihtumista tapahtuu eri ikäluokissa ja sukupuolittain. Yhtenä tavoitteena on myös tuottaa mahdollisuuksia löytää uusia aiheita työkyvyttömyyseläketutkimuksen ja rekisteriaineistojen näkökulmasta. Eläkehakemusten polkujen ja ketjuuntumisen tutkimista valitulla menetelmällä ei ilmeisesti ole aiemmin toteutettu.

Sisällöllisesti luvussa 2 tarkastellaan kansainvälistä ja suomalaista tutkimuskirjallisuutta työkyvystä ja työkyvyttömyyseläkkeistä. Luvussa 3 esitellään teoreettiset viitekehykset, tutkimuskysymykset ja keskeiset käsitteet. Luvussa 4 käydään läpi aineistoa ja menetelmää. Luvussa 5 esitellään tulokset, luvussa 6 käydään läpi pohdintoja ja kriittisiä näkökulmia ja lopuksi luvussa 7 käsitellään yhteenvetoa ja jatkotutkimusta.

## 2 Aiempi tutkimus

Työkyvyttömyyseläkkeiden, eläkkeiden, eläkkeen kaltaisten etuuksien sekä terveyden välisiä yhteyksiä on tutkittu runsaasti ja eri näkökulmista niin kansainvälisesti kuin suomessakin. Eri maissa on erilaiset eläkejärjestelmät, lainsäädännöt ja työkyvyttömyyden määritelmät, joten tutkimustulosten yksiselitteinen yhteen saattaminen on usein hankalaa (Angrist & Pischken 2010). Monissa maissa eläkkeen ja terveyden välisen yhteyden tutkimus koskee pääosin vanhuuseläkeläisiä ja sitä, miten terveys muuttuu, kun työn kuormitus poistuu. Lisäksi monissa maissa ei ole suomalaisen työkyvyttömyyseläkkeen kaltaista eläkettä, vaan erilaisia työkyvyn puutteeseen liittyviä etuuksia. Esimerkiksi Isossa-Britanniassa työkyvyttömät voivat saada tukea useiden etuuksien mukaan riippuen työkyvyttömyyden kestosta, vakavuudesta sekä henkilökohtaisista tarpeista. Eläkkeiden ja eläkettä vastaavien etuuksien määritelmien periaatteissa on paljon eroja, mutta työkyvyttömyyseläkkeiden ja eläkkeenkaltaisten etuuksien perustavanlaatuinen tausta-ajatus on yhtäläinen: henkilön on tarkoitus tehdä työtä, kunnes määritellyn rajan jälkeen katsotaan, ettei hänen tarvitse tai ettei hän voi.

Jotta työkyvyttömyyseläkkeisiin ja työkyvyttömyysetuuksiin liittyvissä tutkimuksissa olisi mahdollista löytää yhteinen konsensus, on tutkimuksia ja empiirisiä analyyseja tarpeen toistaa eri ajanjaksoina ja eri paikoissa. Angristin ja Pischken (2010) mukaan tutkimusten toistaminen, empiirisen aineiston kerääminen ja laajat meta-analyysit ovat tapa päästä kohti yleisempiä tuloksia, joiden kautta on mahdollista tuottaa kulloiseenkin aikaan ja paikkaan tarpeellista tutkimusta. Tämä pätee myös eläkejärjestelmien ja -lajien tutkimuksessa, sillä kenttä on monimutkainen ja työkyvyttömyyseläkkeen sekä vastaavien etuuksien määritelmät vaihtelevat.

Eläketutkimuksen kentällä aiemmat tutkimukset eivät aina ole päätyneet selkeään yhteiseen konsensukseen. Tämä on ohjannut tutkijoita rakentamaan vahvoja identifiointistrategioita, korostaen monipuolisia mekanismeja ja vaikutuksia, jotka pyrkivät selittämään keskimääräisiä vaikutuksia. Näiden toisinaan epätarkkojenkin tutkimustulosten pohjalta syntyy kuitenkin eläketutkimuksen tutkimuskenttä, jota jatketaan, ja jonka tulosten erot on mahdollista selittää sekä menetelmien moninaisuudella että käytetyillä indikaattoreilla. (Garrouste & Perdrix, 2022.) Tämänkaltaisen tutkimuksen reuna-alueille jää väistämättä tilaa tarkentaa

tutkimuskatsetta yksityiskohtiin erityisesti silloin, kun tarkastellaan paikallisen eläkejärjestelmän rajattuja osa-alueita, kuten Suomen julkisen sektorin työkyvyttömyyseläkeratkaisuja viimeisien vuosien aikana.

Tässä luvussa kartoitetaan ensin aiempaa tutkimusta kansainvälisestä näkökulmasta ja sitten suomalaisesta tutkimuksesta käsin. Tarkempi teoreettinen viitekehys määritellään luvussa kolme.

## **2.1 Kansainvälinen tutkimus**

Työkyvyttömyyttä on tutkittu paljon niin Euroopassa kuin laajemminkin. Tämän työn kohderyhmän ja aineiston näkökulmasta on perusteltua tarkastella erityisesti pohjoismaista työkyvyttömyyseläketutkimusta, sillä suomalaisen julkisen sektorin työkyvyttömyyshaasteita on luontevaa kehystää samankaltaisten hyvinvointivaltioiden tutkimuksilla.

Ruotsalaisessa tutkimuksessa havaittiin, että demografiset, sosioekonomiset ja työkyvyttömyysprosessiin liittyvät tekijät vaikuttavat työikäisen väestön työkyvyttömyyseläkkeen hakemisen ja myöntämisen todennäköisyyteen. Alemman sosioekonomisen luokan ammattien todettiin olevan yhteydessä tuki- ja liikuntaelinsairauksiin, jotka voivat estää työtehtävien suorittamisen, jolloin henkilöt hakevat usein kuntoutusohjelmaa tai työkyvyttömyyseläkettä. Lisäksi alempi sosioekonominen asema ja manuaalinen työ vähentävät todennäköisyyttä palata työhön työkyvyttömyyseläkkeen jälkeen. Myös demografisilla tekijöillä on ennustettava rooli työkyvyttömyyseläkkeen hakemisessa ja myöntämisessä: hakemusprosentit kasvavat iän myötä, mutta työkyvyttömyyseläke myönnetään todennäköisimmin alle 30-vuotiaille tai yli 49-vuotiaille. (Perhoniemi ym. 2020.) Nämä tulokset ovat linjassa Norjassa tehdyn työkyvyttömyyseläkkeen hakijoiden sosiaaliturvaa ja työmarkkinoiden taustaa tarkastelevan tutkimuksen kanssa, jossa alempi sosioekonominen asema, manuaalinen työ ja heikompi työhistoria lisäsi työkyvyttömyyseläkkeen hakemista, mutta vähensi eläkkeen myöntämistä (Galaasen ym. 2012).

Tanskalaisessa työkyvyttömyyseläketutkimuksessa havaittiin, että toiminnalliset häiriöt ovat usein yhdistelmä somaattisia ja psyykkisiä oireita, ja ne määritellään yleensä oireyhtymiksi, joita ei voida täysin selittää lääketieteellisillä tai psykiatrisilla

tiloilla. Nämä monimutkaiset oireprofiilit voivat olla ristiriidassa sen kanssa, mitä on kutsuttu ihanteelliseksi työkyvyttömyyseläkkeen saajaksi, eli henkilöksi, jolla on selkeä lääketieteellinen peruste tunnustetusta lääketieteellisestä tilasta, joka on fyysisesti näkyvä, pysyvä ja ei itse aiheutettu. (Mathisen ym. 2021.) Ei ole siis yksiselitteisen selkeää tarkastella työkyvyttömyyden ensisijaisen lääketieteellisen syyn vaihtumista, mikä tulee huomioida tämänkaltaisessa tutkimuksessa.

Ruotsissa on tutkittu eri somaattisten ja psykiatristen sairauksien yhteyttä työkyvyttömyyseläkkeeseen ja kuolleisuuteen 19–60-vuotiaissa pakolaistaustaisten ja Ruotsissa syntyneiden välillä. Tutkimuksessa ei havaittu sukupuolieroja Ruotsissa syntyneiden ja pakolaistaustaisten välillä, mutta pakolaistaustaisuudella oli tilastollisesta merkitsevä yhteys niin somaattisissa häiriöissä (kuten tuki- ja liikuntaelinsairaudet) kuin psykiatristen sairauksienkin osalta työkyvyttömyyseläkkeitä hakiessa. Pakolaistausta altisti kaksinkertaisesta jopa 19-kertaiseen työkyvyttömyysriskiin verrattuna ruotsalaissyntyisiin henkilöihin. Tutkimus osoitti myös, että samanaikaiset mielenterveyden häiriöt selittivät osan suuremmasta työkyvyttömyyseläkkeen riskistä kaikilla henkilöillä, jos taustalla oli somaattisia häiriöitä. (Helgesson, ym. 2021.) Suomessa julkisella sektorilla on runsaasti matalan kynnyksen työpaikkoja, jotka eivät vaadi korkeaa koulutustasoa tai erityisiä ammattitaitoja, ja voivat siten tarjoavat pakolaistaustaisille henkilöille helpomman pääsyn työmarkkinoille.

Vastaavia tuloksia on saatu myös ilman pakolaistausta-muuttujaa. Heikkala ym. (2023) osoittivat, että suuren riskin potilailla, joilla on sekä somaattisia että mielenterveydellisiä häiriöitä, on todettu 7,5-kertainen määrä sairauspoissaoloja kahden vuoden seurannassa, ja heillä on myös 16-kertainen riski jäädä työkyvyttömyyseläkkeelle verrattuna potilaisiin, joilla on vain jompaakumpaa. Samanaikaiset sairaudet tuottavat suuremman työkyvyttömyysasteen verrattuna henkilöihin, joilla on vain yksi todettu työkyvyttömyyden syy.

Ruotsalaisessa työkyvyttömyyseläketutkimuksessa havaittiin klusterianalyysin avulla, että miehet ja naiset eroavat jonkin verran sosioekonomisen aseman suhteen. Tutkimuksen mukaan Ruotsissa suurin syy työkyvyttömyyseläkkeelle jäämiseen on mielenterveyshäiriöt, erityisesti masennus. Vuonna 2008 Ruotsissa tiukennettiin edelleen työkyvyttömyyseläkkeen kelpoisuusvaatimuksia, mikä aiheutti pysyvää työkyvyn heikkenemistä 30-vuotiaille ja sitä vanhemmille. Tutkimuksessa havaittiin

suuremmilla tuloilla olevan yhteys matalampaan työkyvyttömyyseläkkeen riskiin, ja myös ikä, koulutustausta ja ammattiasema vaikuttivat myös merkittävästi työkyvyttömyyseläkkeen saamiseen. (Virtanen ym. 2022.)

Kun yksittäisten tutkimustulosten sijaan tarkastellaan meta-analyysien tuloksia, Garrouste ja Perdrix (2022) osoittivat, ettei aiemmat työkyvyttömyyseläketutkimukset aina päädy yhteensovitettaviin tuloksiin, ja se on johtanut monimutkaisen tutkimuskentän pirstaloitumiseen. Kuitenkin erilaisten soveltuvien lähestymistapojen kautta on mahdollista löytää yhteisiä käsityksiä, teoreettista taustaa sekä menetelmiä myös yksittäisten ilmiöiden kartoittamiseksi. Tutkimustulosten ristiriitaisuutta ja eläkkeiden vertailun hankaluutta lieventää myös se, että aiemmassa tutkimuksessa on osoitettu, ettei eläkkeen määritelmällä sinällään ole merkitystä saaduissa tuloksissa. Valitulla tutkimusmenetelmällä on kuitenkin tärkeä rooli, erityisesti terveystittareiden ja identifiointistrategian osalta, kun pyritään tunnistamaan ja eristämään syy-seuraussuhteita (Nishimura, 2018). Tämän maisterintutkielman osalta ei pyritä selvittämään syy-seuraussuhteita, mutta pyritään tunnistamaan niihin soveltuvia jatkotutkimusaiheita ja -menetelmiä.

Työkyvyttömyyden, eläkkeen ja terveyden väliset yhteydet ovat myös yhteiskuntapoliittinen huolenaihe. Länsimaissa väestö ikääntyy, ja useimmissa Euroopan maissa eläkkeellä olevien määrä kasvaa nopeammin kuin työntekijöiden määrä. Tämä aiheuttaa kestävyysongelmia myös eläkejärjestelmissä, koska ainakin hetkellisesti maksajia on vähemmän kuin edun saajia. Tämä väestödemografinen kehitys on yhtenä keskeisenä syynä johtanut eläkereformien määrän lisääntymiseen Euroopan maissa 1990-luvun lopusta lähtien. Carone ym. (2016) osoittivat, että keskimääräinen vuosittainen eläketoimenpiteiden määrä Euroopassa on moninkertaistunut viimeisten vuosikymmenten aikana, eli noussut yhdeksästä vuosien 1990 ja 1999 välillä aina 44:ään vuosien 2009 ja 2014 välillä. Eikä merkkejä trendin muutoksesta ole toistaiseksi näkyvissä. Suomessa työkyvyttömyyseläkkeet ovat osa eläkejärjestelmää, ja eläkejärjestelmän reformit koskevat siten myös työkyvyttömyyseläkkeitä. Euroopassa on ollut nähtävissä eläkejärjestelmien yhdistämisiä ja lakkauttamisia. Iso osa eläkeuudistuksista on johtanut eläkeiän nostamiseen, ja kun vanhuuseläkkeen alaikäraja on noussut, niin tällöin myös työkyvyttömyyseläkkeen yläikäraja nousee. Esimerkiksi Laaksosen ym. (2022) tulosten mukaan viidenkymmenen ikävuoden jälkeen työkyvyttömyysriski alkaa

kasvaa, joten eläkeiän muutosten väistämättömänä seurauksena Suomessa tulee olemaan entistä enemmän vanhempia työikäisiä työkyvyttömiä.

## 2.2 Suomalainen tutkimus

Myös Suomessa on eurooppalaisen trendin kaltaisesti tehty eläketurvaan merkittäviä suuria muutoksia ja vuosittain pienempiä korjauksia. Suomessa ensimmäinen laaja-alainen sosiaalivakuutus koko väestölle tuli voimaan vuonna 1939, ensimmäiset työkyvyttömyyseläkkeet myönnettiin muutama vuosi myöhemmin vuonna 1942.

Tuolloin työkyvyttömyys ilmeni sodissa haavoittuneiden sotainvalidien kautta, kun he eivät fyysisesti olleet enää työkykyisiä. Sodan psyykkiset kuormitukset jäivät takalalle. Työeläkelaki astui voimaan vuonna 1962. (ETK Eläkejärjestelmä 2024.)

Kansaneläkkeen rakenne ja sen suhde työeläkkeisiin on vuosien varrella muuttunut ja toiminut osin erillään. Yksi merkittävimmistä viimeaikaisista muutoksista työeläketurvan kehityksessä tapahtui vuonna 2005, kun tehtiin laaja eläkeuudistus, joka yhtenäisti työeläkejärjestelmää. Viimeisin suuri eläkeuudistus tuli voimaan vuoden 2017 alussa, ja hallitus ja työmarkkinajärjestöt valmistelevat seuraavaa suurempaa eläkeuudistusta vuosille 2023–2025. (KELA Kelan historia 2024.)

Suomessa työkyvyttömyyseläkettä edeltää käytännössä aina pidempi sairauspoissaolojen jakso. Aiemman tutkimuksen mukaan Suomessa eri ikäluokkien työkyvyttömyyden syyt voivat olla keskenään erilaisia. Ymmärtämällä terveyteen ja hyvinvointiin vaikuttavia tekijöitä jo varhaisessa vaiheessa ja keskivaiheessa urapolkua on mahdollista tunnistaa korkean riskin työkyvyttömyyden ryhmiä. Työkyvyttömyysriskin varhainen havaitseminen ja tähän puuttuminen vähentää sairauspoissaoloja, työkyvyttömyyttä ja voi estää työntekijöiden pysyvän poistumisen työmarkkinoilta. (Mänty ym. 2022.) Ervasti ym. (2019) mukaan vähäisen työn hallinnan ja työkyvyttömyyden yhteys vahvistuu iän myötä, ja työkyvyttömyyden sekä kroonisten sairauksien välinen yhteys on vahvempi nuoremmilla työntekijöillä kuin vanhemmilla. Lisäksi on näyttöä, että puuttuminen sairausperusteisiin poissaoloihin varhaisemmalla iällä voi pienentää riskiä jäädä kokonaan työkyvyttömäksi myös myöhemmällä iällä (Mänty ym. 2022).

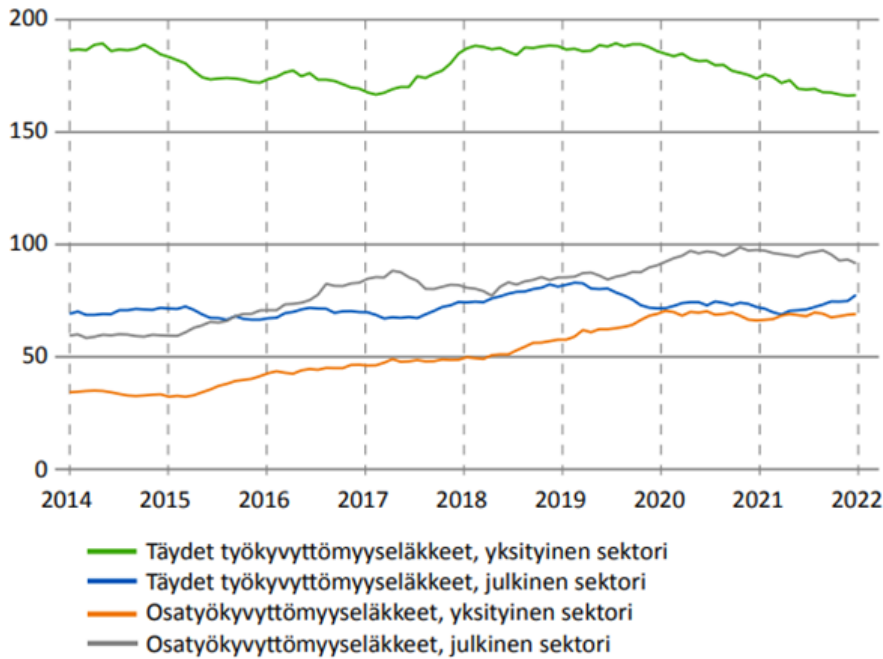
Heinosen ym. (2022) mukaan psykiatrian sairauksista on tullut Suomessa tuki- ja liikuntaelinsairauksien rinnalle yksi suuri työkyvyttömyyden syy, ja trendi näyttää kasvavan huolestuttavaan suuntaan. Yksittäiset suurimmat psykiatrian sairauksista

johtuvat työkyvyttömyyden syyt liittyvät masennukseen ja ahdistuneisuuteen, ja nämä syyt näkyvät sekä miehillä että naisilla. Erityinen huolenaihe tässä väestöryhmässä ovat nuoret aikuiset, joiden sairauspoissaolot ja työkyvyttömyyteen liittyvät poissaolot ovat olleet kasvussa viimeisen kahden vuosikymmenen ajan (Heinonen ym. 2022). Kelan tilastojen mukaan psykiatriset sairaudet ovat jo ohittaneet sairauspoissaolojen perustana tuki- ja liikuntaelinsairauksista johtuvat syyt, ja tämä näkyy erityisesti alle 45-vuotiailla naisilla. Pelkästään vuonna 2023 Kela korvasi noin 5,4 miljoona sairausrahapäivää, joka vastaa noin 20 000 henkilötyövuotta. (KELA Mielenterveyden häiriöt 2024.)

Toisaalta yhä useammat siirtyvät työkyvyttömyyseläkkeelle iän myötä. Laaksosen ym. (2022) mukaan erityisesti 50 ikävuoden jälkeen siirrytään työkyvyttömyyseläkkeelle, koska terveysongelmat yleistyvät ja työkyky saattaa heiketä. Lähempänä vanhuuseläkeikää työkyvyttömyyseläkkeelle siirtyminen voi kuitenkin vähentyä. Tähän saattaa vaikuttaa erilaisten vaihtoehtojen lisääntyminen poistua työmarkkinoilta sekä taloudelliset seikat, kuten työttömyysturvan kannattavuus suhteessa työkyvyttömyyseläkkeeseen. Lisäksi lyhytaikaisissa työkyvyttömyystapauksissa korvaukset voivat tulla sairauspäiväraahajärjestelmästä, eikä työkyvyttömyyseläkkeelle ehkä edes ehditä, jos sairaus ilmenee lähellä vanhuuseläkeikää. Viimeisinä vuosina ennen vanhuuseläkeikää ihmiset saattavat myös sinnitellä työssä, vaikka olisivat oikeutettuja työkyvyttömyyseläkkeeseen. (Kannisto 2022.) Vuoden 2017 eläkeuudistuksen myötä tullut osittainen vanhuuseläke on myös tuonut vaihtoehtoja työuran viimeisille vuosille ja voi vaikuttaa työkyvyttömyyseläkkeen hakemiseen, mutta tästä tarvitaan lisää tutkimusta (ETK Osittainen vanhuuseläke 2024). Näiden yllä mainittujen huomioiden lisäksi alinta eläkeikää on alettu nostaa elinaikakertoimen avulla, jolloin työntekijän aika vanhuuseläkkeen alarajaan kasvaa. Tämä voi lisätä työkyvyttömyyseläkkeen hakemista erityisesti niiden vanhempien henkilöiden osalta, jotka olisivat lähestymässä vanhuuseläkeikää, mutta joiden eläkkeen alaikäraja onkin noussut. Tällöin voi muodostua kokonaan uusia entistä korkeampia ikäluokkia työkyvyttömyyseläkkeiden joukkoon.

Vuoden 2017 työeläkeuudistuksen jälkeen yli 60-vuotiaiden työkyvyttömyyseläkkeelle siirtyminen on lisääntynyt. Erityisesti osatyökyvyttömyyseläkkeiden määrä on kasvanut huomattavasti, kun taas täyden





Kuvio 2. Työkyvyttömyyseläkkeelle vuosina 2014–2021 kuukausittain siirtyneet 60–62-vuotiaat. Nivalainen ym. 2023)

Osatyökyvyttömyyseläkkeelle siirtyneiden määrän kasvu on ollut melko tasaista niin yksityisellä kuin julkisellakin sektorilla vuoden 2015 alusta lähtien, mutta vuodesta 2020 lähtien nousu on tasaantunut. Kuviossa 2 näkyy tasoitettu trendi 60–62-vuotiaiden ikäluokissa työkyvyttömyyseläkkeelle jääneistä vuosien 2014–2021 aikana. (Nivalainen ym. 2023.) Vuoden 2020 jälkeinen vaikutus voi olla pysyvää positiivista kehitystä, mutta se voi olla myös seurausta koronan aiheuttamasta hoitovelasta, joka alkaa realisoitua vasta myöhemmin. Eläketurvakeskuksen mukaan ikäryhmittäin tarkasteltuna vuoden 2020 jälkeen työkyvyttömyyseläkkeiden määrän lasku on ollut tasaista, mutta yli 60-vuotiaiden määrä on pysynyt ennallaan, mitä selittää vanhuuseläkeiän nousu (ETK 2022). Iän vaikutus työkyvyttömyyseläkkeiden alkavuuteen on tilastollisesti merkitsevä, ja hakijan ikä vaikuttaa eri tavoin nuoremmilla ja vanhemmilla (Nivalainen 2023).

Iän lisäksi myös sukupuoli vaikuttaa työkyvyttömyyseläkkeen alkavuuteen. Miehistä 80 prosenttia työskentelee yksityisellä sektorilla, sen sijaan naisista yli puolet työskentelee julkisella sektorilla. Julkisella sektorilla naiset jäävät miehiä useammin eläkkeelle alimmassa eläkeiässä ja naiset myös jäävät hieman miehiä useammin työkyvyttömyyseläkkeelle. Vuosien 2014–2021 välillä naissukupuolen osuus kaikissa työkyvyttömyyslajeissa on ollut 51–55 prosenttia. (Nivalainen ym. 2023.) Tässä työssä käytetyssä tutkimusaineistossa kaikista hakijoista miehiä on noin neljäsosa ja

naisia kolmeneljäsosaa. Naisvaltaisilla aloilla ammattinimikkeet ja työtehtävät ovat hieman erilaisia, ja ammattien yhteyttä sairauspoissaoloihin ja työkyvyttömyyteen on tutkittu paljon. Ammattinimikkeitä ei kuitenkaan käsitellä tämän tutkielman osalta.

Tämä maisterintutkielma jatkaa Kevan aiempaa työkyvyttömyystutkimuksen ja rekisteridatatutkimuksen linjaa, jolle muun muassa Ruponen (2023) on diplomityöllään luonut tilaa ja mahdollistanut tarkentavaa tutkimusta.

### 3 Tutkimuskysymys, käsitteet ja teoreettinen viitekehys

#### 3.1 Tutkimuskysymys

Tässä maisterintutkielmassa tarkastellaan julkisen sektorin työkyvyttömyyseläkkeiden ensisijaisen lääketieteellisen syyn vaihtumista tuki- ja liikuntaelinsairaudesta psykiatrian sairaudeksi työkyvyttömyyden pitkittyessä. Työssä pyritään selvittämään, kuinka paljon julkisella sektorilla esiintyy tautivaihtuvuutta tuki- ja liikuntaelinsairautena (esimerkiksi lanneselän kipu) alkavista työkyvyttömyyseläkkeistä psykiatrisiksi (esimerkiksi keskivaikea masennustila) sairaudeksi silloin, kun työkyvyttömyyseläkeratkaisuja on useampia ja kun työkyvyttömyyden ensisijainen diagnoosi vaihtuu. Lisäksi tarkastellaan, esiintyykö aineistossa eroja taustamuuttujien (ikäryhmä, sukupuoli) puitteissa.

Menetelmänä käytetään sekvenssianalyysia ja klusterianalyysia. Tämä lähestymistapa on eksploratiivinen, ja tarkoituksena on tuottaa uusia havaintoja valitusta aiheesta ilman tarkkarajaisia tutkimushypoteesia. Menetelmä on sosiaalitieteiden kontekstissa hieman vähemmän käytetty, joten on tarkoituksenmukaista kuvata huolellisesti siihen liittyviä valintoja ja yksityiskohtia.

Taustamuuttujina käytetään hakijan ikää ja sukupuolta. Ammattinimikkeet on rajattu pois, vaikka nekin olisivat tärkeä muuttuja tutkittavaksi jatkossa. Tässä yhteydessä ammattinimikkeet kuitenkin kasvattaisivat käsiteltävää materiaalia ja muuttujien määrää turhan suureksi, tuottaisi liikaa informaatiota menetelmän kannalta ja hankaloittaisi tulosten tulkintaa.

Julkiselle sektorille on ominaista naisvaltaisuus, ja henkilöstön ikärakenne on vanhempaa kuin yksityisellä sektorilla. Nuorilla esiintyy hieman eri sairauksia kuin vanhemmilla, mikä näkyy myös käytetyssä rekisteriaineistossa.

Työkyvyttömyyseläkkeitä ja niiden kaltaisia etuuksia on tutkittu monilla tieteenaloilla ja useilla eri käsitteillä. Tutkimuksessa esiintyy toisinaan käsitteellistä epäyhtenäisyyttä, siksi on syytä täsmentää käytettyjä käsitteitä huolellisesti. Seuraavassa alaluvussa esitellään Suomalaisessa kontekstissa keskeisiä käsitteitä tämän työn puitteissa. Erityisesti työkyvyttömyyseläkkeen käsitettä ja ensisijaisen lääketieteellisen diagnoosin määräytymistä on tässä yhteydessä aihetta esitellä seikkaperäisesti.

### 3.2 Käsitteet

**Työkyvyttömyyseläke.** Eläketurvakeskus (ETK työkyvyttömyyseläke 2024) määrittelee työkyvyttömyyseläkkeen seuraavasti: ”Sairauden, vian tai vamman jatkuessa vähintään vuoden ajan työeläkelaitos tai Kela alkaa maksaa kuntoutustukea tai työkyvyttömyyseläkettä. Työkyvyttömyyseläke voidaan myöntää työeläkelaitoksesta 17 vuotta täyttäneelle henkilölle, joka ei ole täyttänyt alinta vanhuuseläkeikää.” Kansaneläkejärjestelmästä työkyvyttömyyseläke voidaan myöntää 16–64-vuotiaalle vakuutetulle. Työkyvyttömyyseläke voidaan myöntää toistaiseksi tai määräajkaisesti, ja se voidaan myöntää täytenä tai osa-aikaisena. Toistaiseksi myönnetty täysi työkyvyttömyyseläke on nimeltään työkyvyttömyyseläke, toistaiseksi myönnettyä osa-aikaista työkyvyttömyyseläkettä kutsutaan osatyökyvyttömyyseläkkeeksi. Määräajkaisesti myönnetty täysi työkyvyttömyyseläke on nimeltään kuntoutustuki ja määräajkaisesti myönnettyä osa-aikaista työkyvyttömyyseläkettä kutsutaan osakuntoutustueksi.

Yleinen työkyvyttömyysmääritelmä edellyttää, että hakijan työkyky on alentunut vähintään 2/5 (40 %) yhtäjaksoisesti vähintään vuoden ajan. Jos työkyky on heikentynyt 3/5 (60 %) tai enemmän, niin eläke myönnetään täytenä työkyvyttömyyseläkkeenä, ja jos työkyky on heikentynyt vähemmän kuin 3/5 mutta vähintään 2/5, niin eläke myönnetään osatyökyvyttömyyseläkkeenä (Taulukko 1).

*Taulukko 1. Työkyvyttömyyseläkkeiden yleinen määritelmä Suomessa*

<b>Eläkelaji</b>	<b>Kesto</b>	<b>Alenema</b>
<b>Työkyvyttömyyseläke</b>	Toistaiseksi	60 %
<b>Osatyökyvyttömyyseläke</b>	Toistaiseksi	40 %
<b>Kuntoutustuki</b>	Määräaikainen	60 %
<b>Osakuntoutustuki</b>	Määräaikainen	40 %

Tätä työkyvyttömyyden jaottelua käyttävät käytännössä kaikki vakuutusyhtiöt Suomessa. Eläketurvakeskuksen (ETK työkyvyttömyyseläke, 2024) mukaan työkyvyn heikentymistä arvioitaessa huomioidaan hakijan jäljellä oleva kyky hankkia itselleen ansiotuloja saatavissa olevalla työllä, jota hänen voidaan kohtuudella edellyttää tekevän huomioiden eläkkeenhakijan koulutus, aikaisempi toiminta, ikä, asuinpaikka

ja muut näihin rinnastettavat tekijät huomioiden. Tämän arvioinnin tekee yleensä eläkelaitoksen vakuutuslääkäri hoitavan lääkärin näkemyksen (b-lausunto) pohjalta.

Käytännössä työkyvyn voi siis menettää joko toistaiseksi tai määräajaksi ja täysin tai osa-aikaisesti. Termillä ”kuntoutustuki” halutaan ylläpitää diskurssia, jossa henkilö on kuntoutumassa. Tavoitteena on, että henkilö palaa kuntoutustuelta töihin tai opiskelemaan. Työkyvyttömyyttä puolestaan pidetään pysyvänä.

**Ikäluokka.** Aiempi tutkimus on osoittanut, että eri ikäluokkien työkyvyttömyyden syyt voivat olla keskenään erilaisia (Mäntä ym. 2022). Työn laadulla on merkitystä työkyvyn muutokseen iän myötä ja työkyvyttömyyden sekä kroonisten sairauksien välinen yhteys on vahvempi nuoremmilla työntekijöillä kuin vanhemmilla (Ervasti ym. 2019). Lisäksi on näyttöä, että puuttuminen sairauspoissaoloihin varhaisemmalla iällä voi auttaa vähentämään sairauspoissaoloja myöhemmällä iällä, ja siten pienentää riskiä jäädä työkyvyttömäksi (Mäntä ym. 2022). Tässä työssä käsitellään kolmea ikäryhmää: alle 45-vuotiaat, 45–60-vuotiaat ja yli 60-vuotiaat. Jaottelu pohjaa yleisesti käytettyyn ryhmittelyyn, mutta perustelu jaolle kolmeen ryhmään on myös menetelmällinen. Sekvenssianalyyseissä pieni määrä luokkia on havainnollinen, eikä tee analyysistä liian monimutkaista tai vaikeasti tulkittavaa. Tutkielmassa ikää on käytetty myös jatkuvana muuttujana aineistoon tutustuttaessa.

**Sukupuoli.** Sukupuoli vaikuttaa itsenäisenä muuttujana, mutta sukupuolella on myös yhteisvaikutusta työkyvyttömyyteen iän kanssa, eikä se siten ole täysin irrallaan iän yhteydestä. Julkisella sektorilla miehet työskentelevät usein erilaisissa ammateissa kuin naiset, ja käyttäytyvät hieman eri tavoin eläkkeen hakemisen suhteen. Lisäksi eri ammateissa on erilaisia mahdollisuuksia jatkaa työtä tai jäädä eläkkeelle. Tässä työssä käytetyssä rekisteriaineistossa on saatavilla tieto hakijan sukupuolesta. Tämä jaoteltu on aiemmalle tutkimusperinteelle tyypillinen binäärinen jako miehiin ja naisiin. Julkinen sektori on naisvaltainen, mutta aineiston koon ansiosta on mahdollista löytää tilastollisesti merkitseviä eroja sekä miehille että naisille.

Ikäryhmiä ja sukupuolia käsitellään tarkemmin luvussa Tulokset ja luvussa Pohdinta.

**Tuki- ja liikuntaelinsairaudet ja psykiatriset sairaudet.** Jako tuki- ja liikuntaelinsairauksiin ja psykiatrian sairauksiin perustuu ICD-10:n luokitukseen (THL ICD-10 2011), joista tässä yhteydessä käytetään selkeyden vuoksi pääosin

nimikkeitä koodien sijaan. Osittain samat tautiluokitukset löytyvät myös ICD-11:stä, mutta ICD-10:n käyttö on perustellumpaa, sillä käytetty rekisteriaineisto kattaa työkyvyttömyyden syyt vuosien 2013–2023 välillä, ja ICD-11:n käyttöönottohanke on alkanut Suomessa vasta vuonna 2023. Sairauksien luokittelu esitellään yksityiskohtaisemmin luvussa neljä.

### **Päädiagnoosin määräytyminen työkyvyttömyyseläkehakemuksessa.**

Fyysisesti pitkäaikaissairailta psykiatrian sairauksien esiintyvyys on 2–3 kertaa suurempi kuin väestössä, ja tuki- ja liikuntaelinsairauksien kivun riski kasvaa psykiatristen sairauksien myötä. Tuki- ja liikuntaelinsairaus tai psykiatrian sairaus kaksinkertaistaa työkyvyttömyysriskin, mutta molempien esiintyminen nelinkertaistaa sen. Esimerkiksi masennus on yleinen pitkäaikaissairauksien liitännäissairaus, mutta sen tunnistaminen voi olla vaikeaa oireiden päällekkäisyyden vuoksi. Hoidot voivat aiheuttaa mielialaoireita tai häiritä hoitomyöntyvyyttä. Yhteisesiintyvyyteen vaikuttavat geneettiset, biologiset, psykologiset, sosiaaliset ja terveyskäyttäytymiseen liittyvät tekijät. Psykiatrian sairaudet heikentävät työkykyä ja työhön paluuta. Tuki- ja liikuntaelinsairauksien ja psykiatrian sairauksien yhteisesiintyvyys lisää merkittävästi työkyvyttömyyden riskiä. (Ryynänen & Metsä 2024.) Samalla tuki- ja liikuntaelinsairauksien esiintyminen psykiatrian sairauksien yhteydessä korostaa sairauksien yhteyttä ja haastaa perinteisen tavan erotella somaattiset sairaudet ja psykiatrian sairaudet. Nykyään monet psykiatrian sairaudet voidaan nähdä yhä useammin yleissairautena, jotka voivat vaikuttaa fyysiseen terveyteen, kun taas somaattiset sairaudet voivat biologisten mekanismien kautta pahentaa psykiatrian sairautta (Gold ym. 2020). Somaattisten sairauksien ja psykiatrian sairauksien yhteiset tekijät, kuten geneettiset tekijät ja ympäristöriskit, voivat lisätä somaattisten sairauksien riskiä ja heikentää terveyttä psykiatrian sairauteen liittyvän haitallisen terveyskäyttäytymisen, kuten epäterveellisen ruokavalion ja tupakanpolton, kautta (Leppävuori 2023). Yhteiset biologiset tekijät viittaavat siihen, että psykiatrian sairauksilla ja somaattisilla sairauksilla voi olla yhteinen tausta. Lisäksi somaattisiin sairauksiin liittyvät kivut ja liikuntakyvyn heikkeneminen voivat pahentaa psykiatrian sairautta, mikä luo itseään vahvistavan kierteen, jossa sairaudet pahentavat toisiaan (Gold ym. 2020).

Monitekijäinen vuorovaikutus ja sairauksien yhteisvaikutus korostavat tarvetta holistiselle lähestymistavalle, jossa useita sairauksia käsitellään samanaikaisesti.

Diagnostiikan haasteet ovat merkittävä tekijä työkyvyttömyyseläkkeiden myöntämisessä, ja sitä kautta tässä työssä käytetyssä aineistossa, sillä työkyvyttömyysratkaisuun merkitään aina vain yksi ensisijainen lääketieteellinen syy. Lääkäreiden tekemät diagnoosit voivat olla ylimalkaisia, potilas saattaa olla haluton kertomaan taustalla olevista psyykkisistä haasteista, tai hän ei välttämättä ole niistä edes itse tietoinen. Myös psykiatrian sairauksien stigma voi vaikuttaa eri ikäluokilla eri tavoin. Lisäksi etävastaanottojen yleistyminen on tuonut mukanaan uudenlaisia haasteita. Tässä tutkielmassa huomioidaan nämä haasteet, vaikka tutkimuskysymyksessä nojataan aineiston ensisijaiseen lääketieteelliseen diagnoosiin. Menetelmä on eksploratiivinen, ja diagnostiikan haasteista huolimatta aineistosta on mahdollista tuottaa uutta tutkimustietoa.

### **3.3 Eläke- ja työkyvyttömyyseläketutkimuksen teorettinen viitekehys**

Tässä luvussa esitellään teorettinen viitekehys eläke- ja työkyvyttömyyseläketutkimuksen näkökulmasta. Ensin käydään läpi teorettisen viitekehysten taustatutkimusta, eläkkeelle siirtymisen ja terveyden yhteyttä, eläkkeelle jäämisen seurauksia sekä työkyvyttömyyteen vaikuttavia yhteyksiä. Tämän jälkeen käydään läpi sekvenssianalyysin teorettista viitekehystä luvussa 3.4, ja vasta luvussa 4 laajennetaan käytettyjä menetelmiä yksityiskohtaisemmin. Jako perustuu työkyvyttömyystutkimuksen ja sekvenssimenetelmän teorettisten viitekehysten erillisyyteen ja erilaisuuteen, sekä valitun menetelmän kuvaustarpeeseen tutkimuskysymyksen kannalta. Työkyvyttömyyden taustateoriaa on osin avattu jo luvussa 2.2, eikä sitä mainittujen osien osalta enää toisteta tässä luvussa.

#### **3.3.1 Teorettiseen viitekehykseen liittyvä taustatutkimus**

Aiemman kirjallisuuden mukaan eläkkeelle jäämisen ja terveysvaikutusten yhteys nähdään usein ristiriitaisina, mutta näitä ristiriitoja on mahdollista sovittaa yhteen. On esimerkiksi tutkittu, mitkä tekijät voivat selittää laajan kirjallisuuden tulosten eroja. Meta-analyysissaan kunkin tutkimuksen tulokset osoittavat, ettei eläkkeen määritelmällä ole merkitystä, mutta valituilla menetelmillä on tärkeä rooli. (Nishimura ym. 2018.)

Aihetta käsittelevissä jatkotutkimuksissa on etsitty konsensusta ristiriitaisten tulosten välille analysoimalla suuri määrä vuoden 2000 jälkeen julkaistuja

vertaisarvioituja akateemisia julkaisuja, jotka käsittelevät eläkkeiden yhteyttä terveyteen (ks. Garrouste ja Perdrix 2022). Analyysin perusteella päädyttiin 51 tutkimukseen kuolleisuudesta, terveydenhuollon kulutuksesta, itse ilmoitetusta terveydestä, masennuksesta, kognitiivisista kyvyistä, fyysisestä terveydestä, patologioista, terveyst indekseistä ja terveyteen liittyvästä käyttäytymisestä. Kyseisessä tutkimuksessa eläkkeitä analysoitiin lisäksi aiemman kirjallisuuden kolmen terveyteen kohdistuvan kausaliteettivaikutuksen valossa: siirtyminen työelämästä eläkkeelle, myöhäisempi eläkkeelle siirtyminen ja aikaisempi eläkkeelle siirtyminen. Tämän työn taustateoriassa nojataan pääosin edellä mainittujen tutkimusten laajoilla meta-analyysillä validoituun teoreettiseen viitekehykseen niiltä osin kuin se liittyy työkyvyttömyyseläkkeisiin.

### 3.3.2 Eläkkeelle siirtymisikä ja terveys

Eläke ja eläkeikä ovat yhdessä sovittuja sosiaalisia konstruktioita, ja ovat yhteydessä esimerkiksi politiikan, kansanterveyden ja elinajanodotteen muutoksiin. Siksi on tärkeä ymmärtää eläkeiän taustalla olevia mekanismeja.

Työmarkkinatutkijoiden meta-analyysin tuloksena on ensinnäkin, että jäädessä eläkkeelle, henkilön terveys usein heikkenee. Jos tämä terveyden heikkenemisen syy olisi voimakas vapaa-ajan lisääntyminen, joka eläköitymiseen liittyy, voisi kuvitella, että progressiivista eläkejärjestelmää olisi suosittava. Näin käsitetään, että työ ja toimeliaisuus ylläpitäisi terveyttä. Toisaalta taas tällainen julkisen politiikan lähestymistapa olisi todennäköisesti epäonnistunut, jos todellinen syy terveyden heikkenemiseen olisikin alentuneet tulot eläkkeellä. On ymmärrettävää, ettei eläköitymiseen liittyviä korrelaatioita ole aina mielekästä pelkistää syy-seuraussuhteiksi. Eläkeikä on mitä suurimmassa määrin resurssikysymys niin eläkkeen saajalle kuin maksajallekin, ja käytössä olevat resurssit ovat yhteydessä terveyteen (Garrouste ja Perdrix 2022). Toisena tutkimuksen aiheena oli analysoida vaikuttaako myöhäisempi eläkkeelle siirtyminen terveyteen (emt.). Eläkepolitiikan näkökulmasta tämä kysymys on tärkeä, koska yhteiskunnallisesta näkökulmasta halutaan yleensä varmistaa, ettei uudistuksilla ole haitallisia vaikutuksia terveyteen. Lisäksi halutaan varmistaa, että tehdyt eläkeuudistukset parantavat julkista taloutta. Eläkepoliittisten uudistusten taustalla on yleensä lopulta kustannusten hillintä, ja eläkeiän nostaminen, joka johtaa pidempään työuriin ja korkeampaan ikääntyneiden

työllisyysasteeseen, on yksi mahdollinen keino tälle. Kolmantena kysymyksenä tutkittiin, vaikuttaako aikaisempi eläkkeelle siirtyminen terveyteen.

Garrousten ja Perdrixin (2022) tulokset osoittivat ensinnäkin, että yleisesti eläkkeelle jääminen tapahtumana johtaa terveydenhuollon kulutuksen vähenemiseen, parempaan itse ilmoitettuun terveyteen, vähäisempään masennukseen, mutta toisaalta kognitiivisten kykyjen laskuun ja kuolleisuuden lisääntymiseen. Lisäksi tuloksena oli, että myöhäisempi eläkkeelle siirtyminen ei vaikuta kuolleisuuteen, mutta vähentää terveydenhuollon kuormitusta eikä vaikuta merkittävästi itse ilmoitettuun terveyteen. Aikaisempi eläkkeelle jääminen voi kuitenkin joissain tilanteista tuoda positiivista vaikutusta itseilmoitettuun hyvinvointiin riippuen sukupuolesta ja ammatillisesta asemasta. Edellä esitetyistä aiheista on toistaiseksi vain vähän tutkimustuloksia, mutta tutkimuksessa (emt.) päädytään siihen, että vaikutusta voi kuitenkin olla juuri naisvaltaisilla aloilla ja sellaisissa ammateissa, joissa fyysinen tai kognitiivinen kuormitus on suurempaa, kuten Suomessa julkisella sektorilla tyypillisesti on.

Eläkkeelle jäämisen myöhentämistä on tutkittu yksityiskohtaisesti myös terveystaloustieteen näkökulmista (esim. Bozio ym. 2021.) Tulosten mukaan yhden vuoden lisäys eläkkeellesiirtymisikään ei luo tilastollisesti merkitsevää positiivista vaikutusta kuolleisuuteen mitattuna 61–79 vuoden iässä. Jos myöhäisemmän eläkkeelle siirtymisen vaikutus kuolleisuuteen olisi havaittavissa, se olisi suuruudeltaan hyvin pieni. Kyseinen tutkimus (vrt. myös Garrouste ja Perdrix 2022) tuottaa kuolleisuuden näkökulmasta positiivista sävyä Suomessakin tehtyihin työeläkeuudistuksiin (mm. elinaikakerroin). On perusteltua kysyä, ovatko samat vaikutukset olemassa myös työkyvyttömyyseläkkeissä, ja tutkia tarkemmin eläkkeelle jäämisen kielteisiä seurauksia. Ainakin elinaikakerroin tuottaa vanhemman joukon työssäkäyviä henkilöitä, joilla on korkeampi työkyvyttömyysriski.

Suomalaisessa tutkimuksessa Laaksonen ym. (2022) toteavat, että psykiatrian sairauksien kanssa sinnitellään työssä useammin kuin somaattisten ongelmien kanssa, ja he kirjoittavat havaintojen viittaavan siihen, että psykiatrian sairaudet eivät aina rajoita työkykyä tai aiheuta poistumista työelämästä yhtä paljon tai yhtä suorasti kuin esimerkiksi tuki- ja liikuntaelinsairaudet. On kuitenkin huomattava, että valittujen terveystilanteiden, muuttujien ja ryhmittelyiden väliset erot voivat vaikuttaa saatujen tuloksien vertailukelpoisuuteen ja itse tuloksiin. Laaksonen ym.

(2022) mukaan muiden terveysongelmien lisääntyessä iän myötä, psykiatrian sairaudet olivat vähiten yleisiä iäkkäämmillä henkilöillä. Tämä näkyi erityisesti 60–64-vuotiaiden ikäryhmässä, jossa työelämän ulkopuolella olevien osuus kasvaa nopeasti. Psykiatrian sairauksien väheneminen iän myötä voi liittyä siihen, että niiden kokeminen on yhteydessä työssäoloon. Nivalainen ym. (2023) lisäävät, että vuoden 2017 jälkeen vanhempiin ikäluokkiin on kohdistunut useita etuusmuutoksia, joiden vuoksi eläkeuudistuksen vaikutusta työkyvyttömyyseläkkeelle siirtymiseen on Suomessa hankala erottaa. Fleischmann ym. (2020) kirjoittavat aiempien tutkimusten osoittaneen, että mielenterveysrasitus ja psykologiset oireet usein lievenevät juuri eläkkeelle siirtymisen jälkeen. Toisaalta voi myös olla, ettei työelämän ulkopuolella samalla tavalla raportoida mielenterveyteen liittyvistä työkyvyttömyyden haasteista, tai ne voivat ikääntyessä jäädä muiden, esimerkiksi tuki- ja liikuntaelimiin liittyvien vaivojen alle. Tästä aiheesta kuitenkin tiedetään vielä hyvin vähän, ja lisää tutkimusta tarvitaan.

### 3.3.3 Eläkkeelle jäämisen negatiiviset seuraukset

Grossmanin (1972) aikaa kestävässä ja usein toistetussa viitekehyksessä esitetään malli, jossa yksilöt kuluttavat terveystalvuita mahdollistaen oman hyvinvointinsa ja työkykynsä. Hyvä terveys parantaa työn tuottavuutta, ja välillisesti lisää hyvinvointia korkeamman kulutuksen kautta. Terveys vaikuttaa myös suoraan hyvinvointiin, sillä ihmiset voivat paremmin ollessaan terveitä. Laajentamalla Grossmanin viitekehystä Galama (2013) tunnistaa kolme mekanismia, jotka selittävät terveyden yhteyttä eläkkeelle jäämiseen riippumatta siitä mikä eläkelaji on kyseessä tai mikä eläkkeelle jäämisen syy on: (1) taloudellisen tilanteen rajoittuminen, eli tulot vähenevät ja mahdollisuus investoida terveyteen tulee suhteessa kalliimmaksi. (2) Muutos terveyshyödyistä saatuun marginaaliin. (3) Oman tehokkuuden ja terveystalvoinneista saadun hyödyn väheneminen.

Galaman (2013) mukaan kun tulot pienenevät, mahdollisuus työterveyspalveluihin poistuu ja terveystalvuiden suhteellinen kustannus kasvaa. Tulojen pieneminen voi vaikuttaa negatiivisesti terveyskäyttäytymiseen, kuten ruoan laatuun ja liikunnan mahdollisuuksiin. Jääminen pois työelämästä vähentää merkityksen tunnetta, joka edelleen vaikuttaa henkilön terveyteen. Tämä vaikutus voi kuitenkin kääntyä myös toisin päin, mikäli lisääntyneellä vapaa-ajalla on mahdollista toimia

merkitykselliseksi kokemalla tavalla (mm. Galama & Kapteyn 2011; Wolfe 1985). Eläkkeelle jäämiseen liittyy usein myös toimintakyvyn alenema. Vaikka tämä alenema ei välttämättä näykään heti eläkkeelle jäämisen seurauksena kaikissa eläkelajeissa, niin tämä toteutuu kuitenkin erityisesti työkyvyttömyyseläkkeissä, jotka ovat määritelmällisesti alenemaa henkilön kyvyssä toimia. Galaman (2013) esittämät kolme elementtiä eivät ole toisiaan poissulkevia, ja ne voivat toteutua yhtä aikaa tai erillisinä. Esimerkiksi kognitiivisen kyvyn heikentyminen saattaa vaikuttaa kykyyn kävellä ja siten heikentää fyysistä terveyttä.

### 3.3.4 Työolojen, kuormituksen, sairauksien ja eläkkeen yhteys työkyvyttömyyteen

Mänty ym. (2015) kirjoittavat työn fyysisen kuormittavuuden lisääntymisen ja työn hallinnan heikentymisen vaikuttavan fyysisen toimintakyvyn laskuun, mutta työolosuhteiden parantuminen voi hidastaa toimintakyvyn heikentymistä. Kouvonen ym. (2017) havaitsivat tutkimuksessaan, että jatkuva altistuminen fyysisesti kuormittavalle työlle oli yhteydessä psyykenlääkkeiden käyttöön. He osoittivat, että kun työn fyysinen kuormittavuus kasvaa, myös psyykenlääkkeiden käyttö lisääntyi. Lisäksi korkeat työn vaatimukset olivat yhteydessä psyykenlääkkeiden käyttöön. Niillä työntekijöillä, joiden altistuminen fyysisesti raskaalle työlle lisääntyi tai oli jatkuvasti korkea, oli suurin osuus psyykenlääkeostoja. Kouvonen ym. (2017) tarkastelivat ostoja myös lääkeryhmittäin, ja havaitsivat, että jatkuva tai lisääntynyt altistuminen fyysisesti kuormittavalle työlle oli yhteydessä uni- ja rauhoittavien lääkkeiden käyttöön. Vastaavasti työn fyysisen kuormittavuuden vähentyminen oli yhteydessä matalampaan masennus- ja rauhoittavien lääkkeiden käyttöön. Aiemmassa tutkimuksessa Kouvonen ym. (2016) osoittivat, että muutokset työn psykososiaalisissa vaatimustekijöissä, kuten työn hallinnan heikentyminen ja voimakkaat työn vaatimukset, olivat yhteydessä mielenterveysoireiden kehittymiseen, ja saattoivat sitä kautta vaikuttaa työkykyyn. Koettu heikompi työn hallinta ja koettu työn vaativuus olivat yhteydessä mielenterveysoireiden kehittymiseen. Aiemman tutkimuksen perusteella voidaan siis sanoa, että työn fyysisen ja psykososiaalisen kuormittavuuden muutokset vaikuttavat tilastollisesti merkittävästi työntekijöiden toimintakyvyn muutoksiin, ja voi edelleen johtaa työkyvyttömyyteen, eikä aina ole selkeää mikä lääketieteellinen diagnoosi alentaa työkykyä. Erityisesti voimakas kuormittavuus ja kuormittavuuden lisääntyminen ovat yhteydessä suurempaan fyysisen toimintakyvyn laskuun ja korkeampaan riskiin

mielenterveysoireiden kehittymiselle. Sen sijaan työolojen parantuminen voi vähentävää toimintakyvyn heikkenemisen riskiä. Näistä tuloksista voi olla hyötyä, kun tunnistetaan työkyvyttömyydessä tapahtuvia muutoksia, ja halutaan etsiä ratkaisuja työkyvyn edistämiseen eri työkyvyttömyyden ryhmissä. Tämä on erityisesti merkityksellistä tämän työn tutkimuskysymyksen näkökulmasta, kun tarkastellaan missä määrin fyysisestä vaivasta aiheutuneet työkyvyttömyydet vaihtuvat psykiatrian sairaudesta johtuvaan.

Ervasti ym. (2018) tutkivat pitkäaikaisen työn fyysistä kuormittavuutta 8–10 vuoden seuranta-aikana. Fyysiseen kuormittavuuteen kuului raskas ruumiillinen ponnistelu tai nosto ja kantaminen sekä selän kiertoliikkeet ja samanlaisena toistuvat liikkeet. Tulokset osoittivat, että raskaan ruumiillisen työn yhteys kuolleisuuteen oli erilainen miehillä ja naisilla. Miehillä jo matalampi kuormitus nosti kuoleman riskiä yli kaksinkertaiseksi verrattuna matalamman kuormituksen ryhmien miehiin. Naisilla vastaavasti kuuluminen korkeimman kuormituksen ryhmään lisäsi ennenaikaisen kuoleman riskiä puolitoistakertaiseksi verrattuna matalampien ryhmien naisiin. Ervastin ym. (2018) tutkimuksen mukaan pitkäaikainen altistuminen hankalille työasennoille oli yhteydessä työkyvyttömyyseläköitymiseen yleisesti ja erityisesti tuki- ja liikuntaelinsairauksien perusteella. Tämä näkyi tuki- ja liikuntaelinsairauksien perusteella myönnettyissä työkyvyttömyyseläkkeissä, jolloin riski oli melkein kolminkertainen verrattuna hankaliin asentoihin altistumattomien ryhmään. Pitkäaikainen altistuminen fyysisesti raskaalle työlle ja hankalille työasennoille kasvattaa tilastollisesti merkitsevästi tuki- ja liikuntaelinperusteisen työkyvyttömyyseläkkeen riskiä. Työkyvyttömyyseläkkeen riski kasvaa jo matalassa altistumisessa, erityisesti silloin kun altistumisen aika on pitkä. Lisäksi altistuminen useammalle hankalalle työasennolle lisää työkyvyttömyyseläköitymisen riskiä. Miesten osalta pitkäaikainen altistuminen raskaalle työlle lisää myös ennenaikaisen kuoleman riskiä.

Työkyvyttömyyttä edeltää sairauspoissaolot, ja sairauspoissaolojen diagnostisilla syillä on yhteys työkyvyttömyyteen. Mänty ym. (2017) havaitsivat, että tuki- ja liikuntaelinsairauksiin, vammoihin, myrkytyksiin ja keuhkosairauksiin liittyvät poissaolot olivat vahvimmin yhteydessä fyysiseen toimintakykyyn ja poissaolot mielenterveyssyistä psyykkiseen toimintakykyyn. Heidän tutkimuksessaan havaittiin, että sairauspoissaolojen määrä oli käänteisesti yhteydessä poissaolojen jälkeiseen

fyysiseen ja psyykkiseen toimintakykyyn, eli mitä enemmän poissaoloja, sitä heikompi toimintakyky oli myös poissaolon päätyttyä. Lisäksi sairauspoissaolojen diagnostiset syyt näkyivät selvästi toimintakyvyssä poissaolojen jälkeen. Tärkeä huomio oli myös se, että sairauspoissaolojen vaikutukset toimintakykyyn voivat olla pitkäaikaisia. Nämä vaikutukset voivat tuottaa pitkäkestoisia sairauksien polkuja, joissa työkyvyttömyyden ensisijaisten diagnoosien merkitys voi jäädä heikommaksi, ja siten vaihtuvuutta ensisijaisissa diagnooseissa voi esiintyä.

Mänty ym. (2018) löysivät pitkittäistutkimuksessaan yhteyden työkyvyttömyyseläkkeen ja työntekijöiden toimintakyvyn välillä. Kun työntekijät siirtyvät vanhuuseläkkeelle heidän fyysinen ja psyykinen toimintakykynsä on yleensä korkeampi verrattuna osa-aikaeläkkeelle tai työkyvyttömyyseläkkeelle siirtyviin työntekijöihin. Fyysisessä toimintakyvyssä ei näkynyt juurikaan muutoksia vanhuus- ja osa-aikaeläkkeelle siirryttäessä, mutta työkyvyttömyyseläkkeelle siirryttäessä tapahtui selkeä alenema. Psyykkisessä toimintakyvyssä näkyi hieman kohoamista vanhuus- ja osa-aikaeläkkeelle siirryttäessä. Tätä kasvua ei ollut näkyvissä työkyvyttömyyseläkkeelle siirryttäessä. (Mänty ym. 2018.) Aiemmassa tutkimuksessa on havaittu, että fyysisesti kuormittava työ ennen vanhuuseläkkeelle siirtymistä aiheuttaa huonompaa fyysistä toimintakykyä ennen ja jälkeen eläkkeelle siirtymisen. Erot vähemmän fyysistä työtä tekeviin näkyivät erityisesti juuri ennen eläkkeelle siirtymisen hetkeä. Kuormittavampaa työtä tekevien toimintakyky hieman parani eläkkeelle jäämisen jälkeen, muttei kuitenkaan samalle tasolle kuin vähemmän kuormittavaa työtä tekevillä. Raskaan työn kuormitus näkyi vielä eläkkeelle siirtymisen jälkeen. (Mänty ym. 2016.) Eläkkeelle siirtyminen vaikuttaa siis tilastollisesti merkittävästi toimintakyvyn muutoksiin. Eläkemuodolla on kuitenkin vaikutusta toimintakyvyn muutoksiin: työkyvyttömyyseläkkeelle siirtyminen on yhteydessä fyysisen toimintakyvyn heikentymiseen, kun taas vanhuuseläkkeelle siirtyminen vaikuttaa positiivisesti psyykkiseen toimintakykyyn. Lisäksi raskas työ ennen eläköitymistä vaikuttaa heikompaan fyysiseen toimintakykyyn sekä ennen että jälkeen eläkkeelle siirtymisen, vaikka toimintakyky hieman parantuisikin eläkkeelle jäämisen jälkeen.

Työn olosuhteet kuten myös työntekijän ikä ja sukupuoli ovat siis eri tavoin yhteydessä työntekijöiden toimintakykyyn ja erilasiin työkyvyttömyyden alttiuksiin. Työn kuormittavuuden lisääntyminen johtaa suurempaan fyysisen toimintakyvyn

heikentymiseen ja korkeampaan mielenterveysoireiden riskiin. Kun huomioidaan aiempi tutkimuskirjallisuus ja nojataan tähän teoreettiseen viitekehukseen, niin on perusteltua tarkastella eksploratiivisilla menetelmillä, missä määrin työkyvyttömyyden ensisijainen lääketieteellinen syy vaihtuu työkyvyttömyyden pitkittyessä tuki- ja liikuntaelinsairauksista psykiatrian sairauksiksi.

Seuraavassa luvussa käydään läpi sekvenssianalyysimenetelmän teoreettista viitekehystä, ja menetelmän tehtyjä valintoja käydään läpi luvussa neljä.

### **3.4 Sekvenssianalyysin teoreettinen viitekehys**

Tutkimusaineiston analysoimiseen käytetään sekvenssianalyysia. Sekvenssianalyysi on eksploratiivinen menetelmä, jota voidaan käyttää sosiaalitieteissä analysoimaan ja ymmärtämään peräkkäisten tapahtumien tai tilojen sarjoja. Menetelmä tarjoaa välineitä tutkia, miten valitut tapahtumat tai tilat seuraavat toisiaan ajan kuluessa. Tämä menetelmä on erityisen hyödyllinen eksploratiivisessa tutkimuksessa kartoittamaan elämäntilanteita, urapolkuja, sosiaalisia prosesseja ja organisaation muutoksia. Päälähteinä sekvenssianalyysin menetelmäosiossa nojataan Gabadinhon ym. (2011) sekä Raabin ja Struffolinon (2023) teoreettisiin viitekehyksiin. Gabadinho ym. (2011) tarkastelevat menetelmän teknisiä yksityiskohtia ja käytännön soveltamista, kuten algoritmien käyttöä ja parametrien säätämistä. Raab ja Struffolino (2023) puolestaan tarkastelevat menetelmien tarjoamaa hierarkkista näkökulmaa sekvenssien välisiin suhteisiin ja menetelmien kykyä ryhmitellä sekvenssejä luonnollisiksi ryhmiksi sosiaalitieteellisessä kontekstissa.

Sekvenssianalyysi ei ole uusi menetelmä. Sitä on käytetty jo 1980-luvulta alkaen, ja sen kehitys on edennyt useassa vaiheessa. Ensimmäisen 10–15 vuoden aikana menetelmää käytettiin enimmäkseen kuvailemaan tapahtumia Optimal Matching (jatkossa OM) -analyyseissa, joissa vertailtiin ja tunnistettiin tapahtumakuvioiden eroja. Näitä tuotoksia käytettiin muuttujina esimerkiksi regressioanalyyseissa. (Abbot & Tsay 2000). Tämä OM-toiminto on edelleen sekvenssianalyysin taustalla, vaikka menetelmä onkin kehittynyt valtavasti viimeisten vuosikymmenien aikana, ja sen hyödyntäminen tutkimuskäytössä on lisääntynyt (Raab & Struffolino 2023, 2). Uusien käyttömahdollisuuksien myötä sekvenssianalyysin käyttö on lisääntynyt moninkertaisesti viime vuosikymmeninä, ja sitä on alettu käyttää enenevässä määrin myös sosiaalitieteissä kuvaamaan yli ajan tapahtuvia muutoksia (Liao ym. 2022).

Lesnard (2009) esittää, että yhteiskuntatieteissä sekvenssianalyysi ja OM-menetelmä mahdollistaa kaksi tapaa tuottaa tutkimusta: joko sekvenssidatan eksploratiivista tutkimista ilman vahvoja ennakko-olettamia tai sellaisten erityisten mallien hakemista, mitä aiempi teoria on ennustanut. Tässä maisterintutkielmassa yhdistetään kumpaakin lähestymistapaa. Aineistosta pyritään tutkimaan minkälaisia kuvioita sekvenssidatasta voi löytyä, mutta tutkimuskysymyksen osalta nojataan aiemman työkyvyttömyyseläketutkimuksen tuloksiin.

Raab ja Struffolino (2023, 1–2) kirjoittavat, että pitkittäistutkimusten myötä kehittynyt sekvenssianalyysi menetelmänä on nykyisellään melko korkealla tasolla ja on hyödynnettävissä monissa tieteissä. Menetelmää käytetään muun muassa lääketieteessä, sosiologiassa, psykologiassa, politiikan tutkimuksessa ja taloustieteessä. Menetelmän taustat ovat tietokonetieteissä, datan louhinnassa ja algoritmimalleissa, joten laskentatehon kasvaessa menetelmän käyttö on mahdollisempaa ja käytännöllisempää kuin aiemmin. Menetelmä sopii erityisesti kuvailevaan tutkimukseen, jossa visualisoidaan sekä tunnistetaan vaiheita ja niiden seurauksia. Liao ym. (2022) esittävät, että sosiaalitieteissä sekvenssianalyysia voidaan käyttää kuvaamaan pitkittäistä elämänkulkudataa, kuten työhistorioita tai perhe-elämän kulkua. Brzinsky-Fayn ym. (2006) mukaan sosiaalitieteen kontekstissa sekvenssi voidaan määritellä esimerkiksi järjestetyksi listaksi kategoristen tilojen (työssä/työttömänä) tai tapahtumien (vanhempien kodista lähteminen, avioliitto, lapsen saaminen) perusteella. Tässä työssä käytetyssä aineistossa kategorioiden tilat ovat työkyvyttömyyseläkeratkaisujen ensisijaiset lääketieteelliset syyt, jotka voivat vaihtua yli ajan (tuki- ja liikuntaelinsairaudet/psykiatrian sairaudet).

Sekvenssianalyysi tutkii, miten erilaiset tapahtumat järjestyvät ja seuraavat toisiaan. Menetelmä keskittyy tapahtumien ajalliseen järjestykseen ja keston merkitykseen. Tämä auttaa ymmärtämään, miten aikaisemmat tapahtumat voivat vaikuttaa myöhempiin. Sekvenssianalyysin avulla voidaan vertailla erilaisia tapahtumaketjuja ja löytää niistä yleisiä malleja tai poikkeavuuksia. Studer ja Rischard (2016) ovat validoineet menetelmän käyttöä yhteiskuntatieteissä omassa tutkimuksessaan, jossa he tekivät vertailevaa tutkimusta monista tavoista mitata tilasekvenssien välisiä eroja. He havaitsivat, että vaikka jotkut etäisyysmittarit eivät toimisi kaikissa tapauksissa, niin ei ole olemassa yleisesti optimaalista etäisyysindeksiä ja että mittauksen valinta riippuu siitä, mihin näkökohtaan tutkimuksessa keskitytään. Yhteiskuntatieteissä

sekvenssianalyysin käyttö eroaa biologian, tietojenkäsittelytieteen ja puhetutkimuksen kaltaisista aloista, siksi tässäkin työssä sekvenssien vertailussa ei tarkastelun kohteena itseasiassa ole yhden sekvenssin muuttaminen toiseksi, vaan se, miten sekvenssit eroavat sosiaalisesti merkityksellisistä näkökohdista.

Tarkasteltaessa sekvenssianalyysin tiloja yhteiskuntatieteiden kontekstissa, on kiinnitettävä erityisesti huomiota tilojen jakautumiseen kussakin sekvenssissä. Voidaan esimerkiksi tutkia, miten työttömänä vietetty aika vaikuttaa henkilön terveydentilaan eläkkeellä. Erot tilojen läsnäolossa tai puuttumisessa tai tilojen jakautumisessa sekvenssien sisällä eivät ota huomioon sitä, miten tilat esiintyvät pitkittäisakselilla. Siksi sekvenssianalyysissä näitä eroja tulisi käyttää yhdessä muiden ulottuvuuksien kanssa. (Studer & Rischard 2016.) Työssä käytetyn aineiston osalta tilojen kestot ovat pääosin rajattuja, sillä työkyvyttömyyden jatkuminen ja uuden ratkaisun saaminen tapahtuvat yleensä rakenteellisesti (työkyvyttömyyttä edeltää sairauspoissaolokausi ja määräaikainen työkyvyttömyyseläkkeen kesto on usein melko samankaltainen). Studerin ja Rischardin (2016) menetelmiä vertailevan tutkimuksen pohjalta tähän työhön sopivimmaksi menetelmäksi käy OM (TRATE-menetelmällä) sekä DHD (Dynamic Hamming). Valinta näiden menetelmien välillä on tämän työn kannalta melko teoreettinen.

Sekvenssimenetelmien määrittelyissä ja kategorioiden matemaattisissa ominaisuuksissa nojataan Elzingan ja Liebroerin (2007) sekä Elzingan (2008) yleisesti käytettyyn teoreettiseen viitekehykseen. Tässä viitekehyksessä sekvenssille aluksi määritellään niin sanottu Alphabet mahdollisten tilojen listana. Sekvenssi on valitun Alphabetin elementin järjestetty lista. Lopuksi sekvensseille määritellään oma aikaleima kullekin tapahtumalle. Tämän jälkeen tyypillisesti käytetään klusterointimenetelmiä ryhmittelemään samanlaisia sekvenssejä yhteen, mikä auttaa tunnistamaan yleisiä polkuja tai trendejä. Lopuksi sekvenssianalyysin tuloksia voidaan visualisoida esimerkiksi sekvenssikaavioina, joka näyttää yksittäisten tapahtumien tai tilojen järjestyksen ja keston.

Seuraavalla luvussa käydään läpi aineistoa ja menetelmää, ja tarkastellaan julkiselta sektorilta työkyvyttömyyseläkeratkaisun saaneiden ihmisten joukkoa. Ensin aineistoa käsitellään aiemman tutkimuksen, teoreettisen viitekehyksen ja tutkimuskysymyksen pohjalta. Tämän jälkeen näkökulmaa laajennetaan Menetelmä-osiossa soveltuvien osin käytetyn sekvenssianalyysin ja klusterianalyysin osalta.

## 4 Aineisto ja menetelmä

### 4.1 Aineisto

Tutkimusaineisto perustuu Kevan rekisteriaineistoon työkyvyttömyyseläkettä hakeneista henkilöistä vuosina 2013–2023. Aineistossa on yhteensä 61 041 henkilöä, joista monilla on useita hakemuksia (esimerkiksi toisen ratkaisun saaneita on 26 816). Aineisto sisältää ratkaisun sekä myönnöstä että hylkäyksestä.

Tutkimusaineistosta poimitaan otoksia, joilla analysoidaan tuki- ja liikuntaelinsairauden syyllä haettuja työkyvyttömyyseläkkeitä, ja tutkitaan yleisellä tasolla missä määrin ensisijainen sairauden syy vaihtuu tuki- ja liikuntaelinsairauksista psykiatrian sairauksiksi sekä miten muutos näkyy valituissa iän ja sukupuolen suhteen.

Tutkimusaineisto on anonymisoitu ja hakijat on järjestetty henkilönumeron mukaisesti. Aineistossa hakijoista on taustatietona sukupuoli, ikä (lääketieteellisen ratkaisun hetkellä), eläkelaji (työkyvyttömyyseläke, osatyökyvyttömyyseläke, kuntoutustuki ja osakuntoutustuki), ratkaisu, ratkaisupäivä, kesto päivinä (edeltävästä ratkaisupäivästä seuraavaan ratkaisupäivään) sekä ensisijainen lääketieteellinen diagnoosi. Aineistossa sukupuoli on koodattu binäärisesti nolaksi ja yhdeksi (0 = mies, 1 = nainen), jolloin keskiarvojen ja muiden tunnuslukujen laskeminen mahdollistuu. Henkilöiden iän perusteella aineistoon on luokiteltu kolme ikäryhmää: Alle 45-vuotiaat, 45–60-vuotiaat ja yli 60-vuotiaat. Tutkimusaineistossa on mukana myös harvinaisemmat eläkelajit: työraueläke sekä kuntoutusnimikkeellä oleva ammatillinen kuntoutus, mutta ne ovat luonteeltaan erilaisia eikä niitä käsitellä tässä työssä tarkemmin. Ratkaisu, ratkaisupäivä ja seuraava työkyvyttömyyseläkkeen laji ovat mukana kulloisessakin tapauksessa. Eläkelajin kesto ilmoitetaan päivinä. Pääasiallinen sairausdiagnoosi perustuu ICD-10 tautiluokitukseen, joka on aineistossa 4 merkin tarkkuudella, esimerkiksi ”F34.1” (Pitkäaikainen masennus). Diagnooseista on summattu pääluokat. Esimerkkinä ICD-10:n koodit F00-F99 ovat ”Mielenterveyden ja käyttäytymisen häiriöt”, ja aineiston ratkaisuilta löytyvät F-alkuiset koodit on koottu ”Psykiatrian sairaudet” käsitteen alle. Johdetut diagnoosien pääluokat näkyvät alla olevasta taulukosta (Taulukko 2). Pääluokat ovat hengityselinsairaudet, hermoston sairaudet, kasvaimet, ruoansulatussairaudet, tuki- ja liikuntaelinsairaudet, verenkiertoelinten sairaudet, muut sairaudet, psykiatrian

sairaudet sekä vammat, myrkytykset yms. (THL ICD-10 2011). Tutkimusaineistossani käytetään näitä yhdeksää pääluokkaa, joista tarkastellaan tuki- ja liikuntaelinsairauksien vaihtumista psykiatrian sairauksiin.

*Taulukko 2. ICD-10 diagnooseista johdetut pääluokat*

<b>Diagnoosien pääluokat</b>		
Hengityselinsairaudet	Muut sairaudet	Tuki- ja liikuntaelinsairaudet
Hermoston sairaudet	Psykiatrian sairaudet	Vammat, myrkytykset yms.
Kasvaimet	Ruuansulatussairaudet	Verenkiertoelinsairaudet

Aineistossa tuki- ja liikuntaelinsairauksien alle kuuluvista diagnooseista tyypillisimmät ovat M50-54 ”muut selkäsairaudet”, M15-17 ”nivelrikot” ja M70-79 ”pehmytkudossairaudet”. Psykiatrian sairauksien alle kuuluvista diagnooseista selvästi useimmin esiintyvä diagnoosiryhmä on F30-39 ”mielialahäiriöt”, seuraavaksi useimmin esiintyvät ryhmät ovat F40-48 ”neuroottiset, stressiin liittyvät ja somatoformiset häiriöt” sekä F20-29 ”skitsofrenia, skitsotyyppinen häiriö ja harhaluuloisuushäiriöt”. Tässä työssä käytetään yllä mainittuja diagnoosiluokkia, eikä eritellä yksittäisten sairauksien tarkempia diagnooseja. Lukuja käydään tarkemmin läpi Tulokset-osiossa.

Tutkimusaineistossa ei ole mukana henkilön ammattitietoa, mutta aiheen jatkotutkimuksissa se olisi hyödyllinen. Julkista sektoria ja naisvaltaista alaa tarkastellessa sukupuoli ja ikä ovat oleellisempia muuttujia valitun tutkimuskysymyksen ja menetelmän kannalta.

Ammattitiedon puutteen lisäksi aineistoon liittyy kaksi muuta tunnistettua haastetta. Ensinnäkin työkyvyttömyyseläkkeen ensisijaisen syyn määrittäminen ei välttämättä aina kaikissa tapauksissa anna todellista kuvaa henkilön tilanteesta. Vaikka diagnoosin tekee viimekädessä aina hoitava lääkäri ja ratkaisu perustuu vakuutuslääkärin arvioon, voi hakijalla olla stigmaan, elämäntilanteeseensa, työuraansa tai muuhun liittyviä syitä hakea muulla kuin ilmeisimmällä lääketieteellisellä syyllä valitsemaansa eläkelajia. Toiseksi työkyvyttömyyteen liittyvissä tapauksissa esiintyy usein yhteisvaikutusta, eli henkilö kärsii samanaikaisesti useammasta kuin yhdestä sairaudesta, kuitenkin työkyvyttömyyseläkeratkaisuun kirjataan aina vain yksi ensisijainen lääketieteellinen

syy. Yksin tämän perusteella ei kaikissa tapauksissa voida päätellä kuinka paljon ensisijainen diagnoosi vaikuttaa työkyvyn alenemaan, vai onko taustalla kenties muita vaikuttavia diagnooseja. Diagnoosit määritellään siten, että ensimmäinen diagnoosi on se, jonka vakuutuslääkäri on arvioinut olevan pääasiallinen työkyvyttömyyden syy.

Tässä maisterintutkielmassa tarkastellaan ensisijaista työkyvyttömyyseläkehakemuksen syytä, mikä on siis yksinkertaistavaa, mutta vastaa alan vallitsevaa tutkimustapaa ja tilastointia. Ensimmäistä diagnoosia käytetään myös sairauksia koskevissa tilastoissa, joita Eläketurvakeskus ylläpitää ja julkaisee (Ruponen 2023). Henkilön koko sairaustaustaa, kuten sairauspoissaoloja, ei tässä työssä tarkastella, vaan keskitytään tutkimuskysymyksen näkökulmasta rekisteriaineiston perusteella esiin nouseviin ilmiöihin. Käytetty menetelmä on eksploratiivinen, ja sen tarkoitus on tuottaa uusia havaintoja ja käyttötarkoituksia tutkimukseen avoimesti ja joustavasti ilman tarkasti määriteltyjä hypoteeseja tai odotuksia.

## **4.2 Menetelmä**

Tässä luvussa esitellään soveltuvin osin käytetyn sekvenssianalyysin menetelmiä. Sekvenssianalyysia käytetään lukuisin eri tavoin monenlaisessa tutkimuksessa, eikä kaikkien eri menetelmien läpikäyminen ole mielekästä tässä yhteydessä. Tässä luvussa esitellään osa sekvenssianalyysin teknisestä soveltuvuudesta käytetyn tutkimuskysymyksen ja aineiston kannalta.

### **4.2.1 Sekvenssianalyysi**

Tyypillinen sekvenssianalyysi sisältää yksinkertaisimmillaan kolme vaihetta: kertomusten tai prosessien koodaaminen sekvensseiksi, sekvenssien välisten eroavaisuuksien mittaaminen vertaillen ja jonkinlainen aineiston rajaaminen, esimerkiksi klusterianalyysi. Gabadinho ym. (2009) mukaan R-ohjelmistolle tehty TraMineR-kirjasto osoittautuu hyödylliseksi työkaluksi näiden eri kategorioiden välisessä muuntamisessa ja sekvenssianalyysin käytössä yhteiskuntatieteellisessä tutkimuksessa.

Sekvenssien esittäminen datatiedostoissa voi vaihdella riippuen siitä, miten aineiston on kerätty, mikä sen laatu on ja miten informaatio on järjestetty. Useinkaan

sekvenssit eivät löydy sellaisenaan aineistosta, vaan ne täytyy muodostaa alkuperäisistä kerätyistä ajanjaksoista. (Gabadinho ym. 2009.) Sosiaalitieteissä tapahtuman seuraava tila voi riippua edellisestä ja siten muista tapahtumista, joita on jo tapahtunut, joten tila- ja tapahtumasekvenssien muuntaminen toiseksi ei aina ole suoraviivaista. Työkyvyttömyyseläkeratkaisujen ensisijaisia syitä tarkastellessa on oltava kriittinen sen suhteen mikä tila edeltää toista, sillä syyt eivät aina ole yksiselitteisiä.

Tapahtumien elinkaarilla on myös keskeinen merkitys sekvenssidatassa, koska ajankohta, kesto ja muutoskohdat ovat tarkastelun kohteena. Gabadinho ym. (2009) kirjoittavat, että tilojen sekvenssien tapauksissa on tärkeää tietää, tapahtuuko tilojen kohdistaminen sisäisen aikaviitepisteen mukaan (esimerkiksi yksilön ikä) vai ulkoisen aikaviitepisteen mukaan (esimerkiksi tammikuusta joulukuuhun). Tässä maisterintutkielmassa käytetyn aineiston kannalta on mahdollista käyttää kumpaa tahansa aikaviitepistettä. Rohwer ja Pötter (2002) esittävät erään diskreetin aikajanatypologian, jolla sekvenssien tilat voidaan määritellä. He erottelevat kalenteriajan ja prosessiajan akselin. Kalenteriajan akselilla ei ole luonnollista alkua, vaan ajan pisteiden määrittäminen on yksinkertaisesti sovitut ajanhetket.

Prosessiajan akselilla alkuperä edustaa aloitustapahtuman päivämäärää, josta eteenpäin prosessia seurataan aina tiettyyn päättymistapahtumaan saakka. Raab ja Struffolino (2023, 15) kirjoittavat prosessiajan sopivan erityisesti sosiaalitieteissä tapahtumien alkuihin ja loppuihin, kuten työkyvyttömyyseläkkeisiin, jossa yksi aikakausi päättyy. Käytettävissä oleva aineisto usein määrittelee, minkälaisissa ajanjaksoissa sekvenssidataa tarkastellaan, eikä ole suurta vaikutusta mitataanko prosessiaikaa päivissä, kuukausissa tai vuosissa.

Työkyvyttömyyseläkeratkaisujen aineiston näkökulmasta on tarkoituksenmukaista käyttää sekvenssianalyysissa prosessiakselilla sovittuna alkupisteenä hakemuksen ratkaisuntekopäivää (rpvm1:stä rpvm15:een), joka on ilmoitettu muodossa VVVV-KK-PP (esimerkiksi ”2014-08-22”). Tutkittaessa pitkittyneitä työkyvyttömyyseläkeratkaisuja loppupisteenä on seuraavan ratkaisuntekopäivän alkupiste riippumatta siitä, onko ratkaisu hylkäys vai myöntö. Sekvenssien kesto on ilmoitettu päivinä, joka näkyy ratkaisujen välisenä aikana. Sisäistä aikaviitepistettä (ikä) käytän muuttujana varsinaisen sekvenssianalyysin ulkopuolella.

Aineiston käsittelyssä sekvenssitaulukoksi käytetään tila-sekvenssi -muotoa (state-sequence, STS). Tämä on TraMineR-ohjelmiston formaatti, jossa peräkkäiset tilat annetaan peräkkäisissä sarakkeissa. Tällöin oletetaan jokaisen sarakkeen vastaavan ennalta määrättyä aikayksikköä (rpvm:t). Sekvenssille määritellään aluksi Alphabet mahdollisten tilojen listana, määritellään aikaleima ja lopuksi ryhmitellään samankaltaisten sekvenssien ryhmiksi (klusterointi). Tämän jälkeen sekvenssianalyysin tuloksia voidaan visualisoida esimerkiksi dendrogrammina ja sekvenssikaavioina hakemuspoluista. Dendrogrammi näyttää hierarkkisen klusterointimenetelmän tuottama puumaisen diagrammin, joka visualisoi ryhmien yhdistymisprosessin ja mahdollistaa ryhmien tarkastelun. Sekvenssikaavio hakemuspoluista kuvaa ajan myötä tapahtuvien tilojen järjestyksen ja muutokset sekvenssianalyysissä.

#### 4.2.2 Sekvenssin Alphabet

Alphabetit ovat kategorisia tiloja, joista sekvenssit muodostuvat. Alphabetit määrittyvät käytännössä aina tutkijan aineiston perusteella, ja tutkimuskysymyksestä riippuu ovatko Alphabetit työn kannalta oleellisia vai ei. On myös tutkimusasetelmia, joissa tällaista päätöstä ole mahdollista tehdä aineiston takia, ja sekvenssianalyysi on tehtävä sen mukaisesti. (Raab & Struffolino 2019, 13). Alphabet voidaan kirjoittaa pilkulla eroteltuna. Esimerkiksi tässä maisterintutkielmassa käytetyssä aineistossa yhden sekvenssin Alphabet voisi olla "x = T, P, M, NA", jossa T on tuki- ja liikuntaelinsairaudet, P on psykiatrian sairaudet, M on muut sairaudet ja NA on polun päättyminen (ei uutta ratkaisua).

Tarkempi Alphabetin määrittely ei välttämättä tuota tarkempaa tulosta.

Sekvenssianalyysissä tarkoituksena on löytää sopiva tasapaino ja yksityiskohtien laatu. Esimerkiksi Alphabet "x = T, T, T, P, P, M, NA" ei välttämättä tuo mitään lisää verrattuna Alphabet "x = T, P, M, NA", mutta ensimmäinen tuottaisi ylimääräistä informaatiota tulosten tulkintaan. Sekvenssianalyysin tarkoitus on vähentää monimutkaisuutta, ja yksinkertaisempi Alphabet on myös visualisoinnin kannalta selkeämpi. Alphabetin tiloja on mahdollista kytkeä yhteen, mutta kytkentä täytyy perustella, koska tällöin osa informaatiosta häviää. Yleisesti parhaan lopputuloksen kannalta on perusteltua aloittaa isommasta ja kulkea iteratiivisesti kohti pienempää ja selkeämpää Alphabetiä. Tätä tarkentamista voi jatkaa OM-menetelmällä ja

klusterianalyysillä, ja vaikutusta voi mitata korrelaatiolla matriisien välillä joko Mantel coefficient -testillä tai permutaatiotestillä. (Raab & Struffolino, 13–14). Yksinkertaisuudessaan sekvenssin Alphabet on ajallisesti järjestetty jono kategorisia elementtejä (Taulukko 3).

Taulukko 3. Esimerkkipoimintoja sekvenssidatasta: Alphabet  $x = (x_1, x_2, x_3, x_4, x_5, x_6)$

Koodi	Konversio	ID	Sukupuoli	Ikäryhmä	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$
STS	Mistä/mihin	81	Mies	45–60	Tuki- ja liik. sair.	Tuki- ja liik. sair.	Tuki- ja liik. sair.			
STS	Mistä/mihin	112	Nainen	45–60	Tuki- ja liik. sair.	Tuki- ja liik. sair.	Tuki- ja liik. sair.	Psykiatrian sair.	Tuki- ja liik. sair.	Psykiatrian sair.
STS	Mistä/mihin	126	Nainen	45–60	Tuki- ja liik. sair.	Tuki- ja liik. sair.	Tuki- ja liik. sair.			

Sekvenssianalyysissä on myös monia muita tapoja esittää Alphabet. Gbadinho ym. (2011) esittelevät STS-menetelmän lisäksi kolme muuta tapaa: SPS, DSS ja SPELL. Näiden menetelmien käyttäminen on tarpeen erityisesti silloin, jos data on monimutkaisempaa, ajallisesti haastavampaa tai lukumäärältään suurempi. Kussakin menetelmässä on omat hyötynsä ja haasteensa, ja menetelmä tulee sovittaa kulloiseenkin tutkimuskysymykseen ja käytettävissä olevaan dataan. Tämän tutkielman aineistoin kannalta STS:n käyttäminen on perusteltua.

#### 4.2.3 Puuttuvat arvot

Sekvenssianalyysissä on tavallista, että puuttuvat arvot korvataan asettamalla alphabetiin sopiva arvo (esimerkiksi "ei sairautta") tai poistamalla tapaukset, joissa puuttuvia arvoja on (esimerkiksi "N/A" tai "Tyhjä"). (Raab & Struffolino 2023, 17–18). Puuttuvien arvojen käsittelylle on itse sekvenssianalyysissäkin omat menetelmänsä. Halpin (2016a) esittää, että puuttuvat arvot voidaan käsitellä ei-itse-identtisinä arvoina, jolloin sekvenssit, jotka sisältävä puuttuvia arvoja ajatellaan keskenään erilaisiksi. Tämä tuottaa paremman tuloksen kuin se, että sekvenssit ajateltaisiin samankaltaisiksi niiden arvojen kanssa, joissa ei ole puuttuvia arvoja. Puuttuvia arvoja on myös mahdollista korjata matemaattisen teorian kautta (esimerkiksi Halpin 2016b). Raab ja Struffolino (2023, 19–20) lisäävät, että yksi tyypillinen tapa käsitellä puuttuvia arvoja on käsitellä dataa jo analysointivaiheessa sellaiseksi, ettei puuttuvista arvoista synny välejä. Tällöin ei aiheuteta ongelmaa tulosten kannalta, koska oletetaan, että todelliset arvot puuttuvien arvojen ympärillä

on tiedossa. Sekvenssianalyysin puuttuvat arvot ovat tunnistettu haaste, ja jatkossa tähän haasteeseen on tehtävä tarkentavaa tutkimusta (Raab & Struffolino 2023, 20).

Tässä työssä käytetyn aineiston osalta puuttuvat arvot eivät ole haaste.

Tarkasteltaessa rekisteriaineistoa, jossa jokaiselle tapahtumalle on olemassa ratkaisu (myöntö tai hylkäys), diagnoosiluokka (esimerkiksi tuki- ja liikuntaelinsairaus tai psykiatrian sairaus) ja ratkaisupäivämäärä, ei puuttuvista arvoista tule ongelmaa.

Käytetyn aineiston ja tutkimuskysymyksen osalta Raabin ja Struffolinon (2023, 19) tapa osoittautuu hyödylliseksi, koska aineiston mukaisessa sekvenssianalyysissa puuttuva arvo on indikaattori työkyvyttömyyseläkeratkaisun hakemuspolun päättymiselle, ja nämä päättymiset on oleellista pitää mukana analyysissa. Tällä tavalla puuttuva arvo käsitetään joka tapauksessa erilaisena kuin sekvensseissä näkyvät diagnoosien pääluokat, eikä ole merkitystä ovatko puuttuvat arvo keskenään identtisiä vai ei.

#### 4.2.4 Optimal Matching -menetelmä

Optimal Matching (OM) -menetelmä on yksi keskeisistä sekvenssianalyysimenetelmistä. OM:ssä kahden sekvenssin erilaisuuden aste määräytyy sen mukaan, kuinka monta painotettua muokkaustoimintoa tarvitaan kahden sekvenssin sovittamiseksi yhteen. Yleensä käytetään kolmenlaisia muokkaustoimintoja: lisäys, poisto ja korvaaminen. OM:n tulos on kaikkien sekvenssien välinen matriisi, joka on yhdistettävä klusterianalyysiin, moniulotteiseen skaalaukseen tai muuhun tietojen vähentämismenetelmään, joka käsittelee erilaisuusobjekteja. (Lesnard 2009.)

Menetelmänä OM on hyödyllinen vertaillessa erilaisia tapahtumasekvenssejä, kuten urapolkuja, elämänvaiheita tai tapahtumaketjuja. OM-menetelmän muokkaustoimintojen avulla voidaan muuntaa yksi sekvenssi toiseksi mahdollisimman pienin kustannuksin, mikä antaa mittarin sekvenssien väliselle etäisyydelle. (Raab & Struffolino 2023.) OM-menetelmän etuna on kyky käsitellä monimutkaisia ja epäsäännöllisiä sekvenssejä, jotka eivät välttämättä ole samanpituisia tai joissa voi esiintyä monenlaisia tilasiirtymiä. Tämä tekee siitä joustavan ja soveltuvan erityisesti sosiaalitieteisiin, joissa analysoitavat sekvenssit voivat olla hyvin monimuotoisia. OM-menetelmän avulla voidaan ottaa huomioon esimerkiksi urapolut, jotka voivat sisältää erilaisia työsuhteiden kestoja, katkoksia ja

uramuutoksia. (Gabadinho ym. 2011.) Saman tavalla tässä työssä vertaillaan työkyvyttömyyseläkeratkaisujen hakemuspolkuja, joiden kestot ovat eri mittaisia ja voivat sisältää katkoksia, sekä tarkastellaan hakemuspolkujen muuntumista tuki- ja liikuntaelinsairauksista psykiatrian sairauksiksi.

#### 4.2.5 Substituutiomatriisi ja TRATE

Gabadinho ym. (2011) kirjoittavat, että substituutiomatriisin käyttö voi parantaa sekvenssianalyysin laatua ja tarkkuutta. Substituutiomatriisin TRATE-menetelmä ottaa huomioon aineiston sisäiset siirtymät ja mukauttaa siirtymien kustannukset niiden yleisyyden mukaan, mikä tekee analyysistä realistisemmän ja luotettavamman. Kustannuksilla tarkoitetaan sitä laskentaa ja resurssia, jota OM-menetelmä käyttää muuntaessaan yhden sekvenssin toiseksi. Kustannusasetannalla on vaikutuksia siihen, millaisia ajallisia malleja OM voi paljastaa, kun sitä sovelletaan yhteiskuntatieteellisiin tietoihin (Lesnard 2009). TRATE-menetelmä on erityisen hyödyllinen monimutkaisissa aineistoissa, joissa tilasiirtymien yleisyys vaihtelee.

Substituutiomatriisin käyttö on perusteltua, koska se mahdollistaa tarkemman sekvenssianalyysin ja parantaa analyysin tulosten luotettavuutta ja tulkittavuutta. Aineiston kannalta TRATE-menetelmä on hyödyllinen, sillä se perustuu sekvenssien transitiomatriisiin, joka ottaa huomioon sekvenssien sisäiset siirtymät tilasta toiseen. Toisin kuin yksinkertaisempi CONSTANT-menetelmä, joka asettaa kiinteän kustannuksen kaikille tilanvaihdolle, TRATE-menetelmä mukauttaa kustannukset tilasiirtymien yleisyyden mukaan. Gabandinho ym. (2011) mukaan tämä menetelmä mahdollistaa tarkemman ja kontekstisidonnaisemman sekvenssianalyysin, ja sopii siten käytetyn aineiston kaltaiseen dataan.

Tämän maisterintutkielman kannalta substituutiomatriisin käyttäminen TRATE-menetelmällä on oleellista sekvenssianalyysissä OM-menetelmän yhteydessä, kun tutkitaan työkyvyttömyyseläkehakemusten ensisijaisen lääketieteellisen syyn vaihtumista tuki- ja liikuntaelinsairaudesta psykiatrian sairaudeksi yli ajan. Menetelmä auttaa saamaan realistisen kuvan sekvenssien välisistä eroista ja varmistaa analyysin yhdenmukaisuuden ja vertailtavuuden. Lähestymistapa huomioi kuinka usein ja millä tavoin hakemusten syyt muuttuvat ajan myötä, mikä puolestaan parantaa analyysin tarkkuutta ja luotettavuutta. (Gabadinho ym. 2011.)

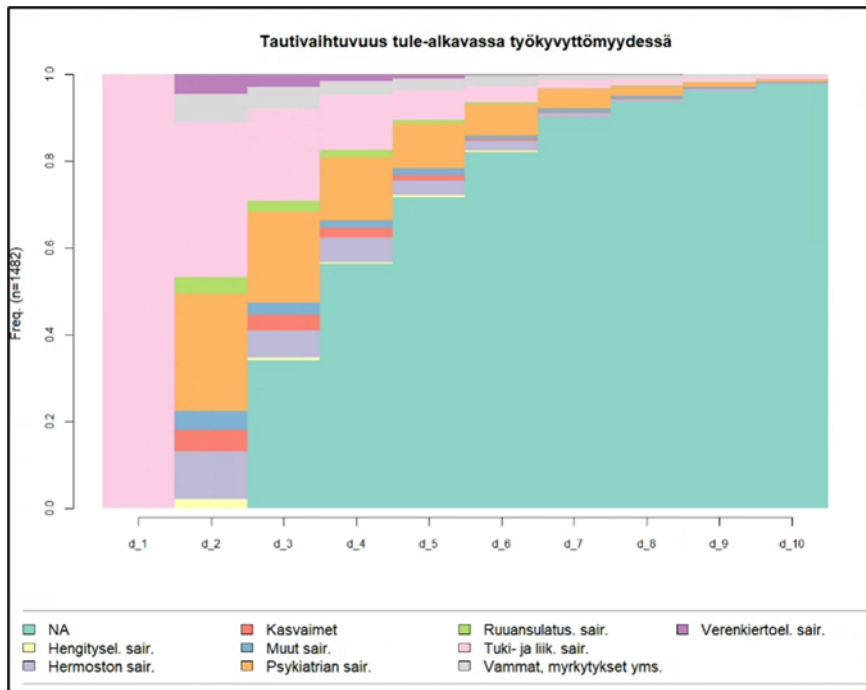
#### 4.2.6 Visualisointi

Sekvenssien esittäminen taulukkona tuottaa valtavasti informaatiota ja tulkintojen esittäminen on haastavaa. Sekvenssejä ei ole mielekästä esittää keskiarvojen tai hajontojen kautta yksittäisinä numeerisin indikaattoreina, kuten esimerkiksi tuloja olisi. Raab ja Struffolino (2023, 26–27) esittävät, että sekvenssin perusteellinen kuvaus edellyttää useiden kategoristen tilojen jakauman selvittämistä sekvenssin eri kohdissa sekä sekvenssikohtaisten yhdistelmien tarkastelua. Pelkän taulukotietojen tarkastelussa olisi riskinä tuottaa niin suuri määrä lukuja, että tarkkuudestaan huolimatta tarpeellisten havaintojen tunnistaminen kävisi mahdottomaksi. Visualisoinnilla on ollut korostuneen tärkeä rooli sekvenssianalyysitutkimuksessa, erityisesti siitä alkaen, kun Stata ja R-ohjelmisto ovat tulleet laajempaan käyttöön, sekä viimeaikaisten teknologisten ja matemaattisten edistymisten myötä, kun aineistojen suuremmat koot ovat mahdollistuneet (esim. Fasang & Liao 2014; Brzinsky-Fay 2014; Raab & Struffolino 2023, 26–50).

Tyypillinen sekvenssien visualisointi vaatii kolmen ulottuvuuden esittämisen: havaintoyksiköt, aikapisteet ja tilat. Nämä kolme yhdessä muodostavat suuren määrän informaatiota, jonka esittäminen selkeästi on haastavaa. Useiden yksittäisten havaintosekvenssien näyttäminen samankaltaisuuksien tai säännönmukaisuuksien löytämiseksi tuottaa vasemmalta oikealle luettavan kaavion, joka koostuu monista päällekkäisistä viivoista. Nämä havaintosekvenssit pyritään yhdistämään niin, että säännönmukaisuudet tulevat näkyviksi. Sekvenssianalyysikaaviot tuottavat usein niin suuren määrän informaatiota, että ne ylikerrostuvat, siksi sekvenssianalyysikaavio on hyödyllistä piirtää koko aineiston sijaan vain valitun otoksen havainnoista. Lisäksi jos datassa on jo olemassa yksiköiden järjestys, sekvenssit yhdessä kaaviossa tulisi järjestää niiden mukaisesti. Myös väreihin on syytä kiinnittää huomiota, sillä kaavioita voidaan käyttää eri julkaisuissa ja on oleellista huolehtia saavutettavuudesta, luettavuudesta ja kustannuksista värien osalta. (Liao ym. 2022.)

Alla esitetään esimerkkikuvio sekvenssidatan visualisoinnista (Kuvio 3, N=1 482). Kyseessä on aineistosta otos, jossa on mukana vain jatkuvat tai vaihtuvat työkyvyttömyydet tuki- ja liikuntaelinsairauksina alkaneista työkyvyttömyyseläkehakemuksista. Kuviossa nähdään vaaleanpunaisella tuki- ja liikuntaelinsairauksina alkaneet tapaukset ja voidaan havaita missä määrin ne

muuttuvat esimerkiksi oransseiksi psykiatrian sairaustapauksiksi. ”N/A” tarkoittaa, ettei uutta ratkaisua ole tehty.



Kuva 3. Esimerkkikuvio aineistosta. Otoksen tautivaihtuvuus tule-alkavassa työkyvyttömyyseläkehakemuksessa

Visualisoinnissa hyödynnän Brzinsky-Fayn (2014) klassista sekvenssikaaviotyyppeä: sekvenssi-indeksikaaviota. Tässä sarakkeet ovat polkutyyppejä, jotka johtuvat klusterianalyysistä. Vaikka yksittäisissä sekvensseissä osoittautuisikin olevan jonkin verran vaihtelua, niin klusterointi tiivistää ne yhteen laskettujen samankaltaisuuksien kautta.

Lopulta suurin haaste sekvenssidatan visualisoinnissa on päätös siitä, mitä tietoa näytetään ja mitä tietoa jätetään näyttämättä. Tämä on jossain määrin ominaista kaikelle empiiriselle tutkimukselle, ja tätä tutkimusaihetta onkin mahdollista laajentaa myöhemmissä tutkimuksissa tämän tutkielman rajausten ulkopuolelle.

#### 4.2.7 Objektit ja ryhmittely (klusterointi)

Kun sekvensoinnin etäisyysmatriisi on laskettu, siirrytään ryhmittelyyn. Raab ja Struffolino (2023, 83) kuvaavat klusterianalyysin menetelmänä, jonka tavoitteena on jakaa data ryhmiin niin, että samassa ryhmässä olevat objektit ovat samankaltaisempia kuin muihin ryhmiin kuuluvat objektit. Sekvenssianalyysin kontekstissa samassa ryhmässä olevien sekvenssien tulisi sisältää jotain peräkkäistä

säännönmukaisuutta, esimerkiksi tilojen peräkkäisyyden tai jakson keston mukaan. Klusterianalyysissa löydettyä typologiaa voidaan käyttää joko kategorisena riippumattomana muuttujana tai kategorisena riippuvana muuttujana regressioanalyysissä. Klusterianalyysi on yleisimmin käytetty tapa tunnistaa ryhmiä, jotka edustavat tietyn ajallisen sosiaalisen prosessin empiirisiä toteutumisia sekvenssitiedoissa. Tässä työssä ryhmittely perustuu diagnoosiluokkiin (tilat) ja ratkaisujen määrään.

Aineiston hierarkkiseen ryhmittelyyn käytetään Agglomerative Nesting (AGNES) -menetelmää, joka yhdistää ”bottom-up” -periaatteella aluksi jokaisen sekvenssin omaksi ryhmäkseen ja sitten toistuvasti yhdistää lähimmät ryhmät, kunnes kaikki sekvenssit ovat yhdessä ryhmässä. AGNES-menetelmän sisällä käytetään Wardin WARD.D2-menetelmää, joka minimoi ryhmien sisäisen varianssin yhdistämisen aikana. Tämä yhdistelmä varmistaa, että samankaltaiset sekvenssit ryhmitellään tehokkaasti yhteen, mikä edelleen auttaa tunnistamaan työkyvyttömyyseläkeratkaisujen lääketieteellisten syiden vaihtumista.

AGNES on yksi suosituimmista hierarkkisista klusterianalyysimenetelmistä, ja sitä käytetään, kun halutaan tarkastella havaintojen välisiä suhteita eri tarkkuustasoilla (Gabadinho 2011). AGNES-menetelmä yhdistää ryhmiä askel askeleelta perustuen niiden välisiin etäisyyksiin. Tämä prosessi jatkuu, kunnes kaikki havaintoyksiköt ovat yhdessä ryhmässä. Tämän hierarkkisen lähestymistavan etuna on, että se tuottaa dendrogrammin, joka visualisoi ryhmien yhdistymisprosessin ja mahdollistaa ryhmien tarkastelun eri aggregaatiotasoilla (täysin erillisistä täysin yhtenäiseksi ryhmäksi). Tulokset-luvussa esitellään ryhmittelyn tulokset sekä dendrogrammi menetelmän havainnollistamisen vuoksi.

AGNES-menetelmä mahdollistaa sekvenssien luonnollisten ryhmittymien tarkastelun ja tunnistamisen. Menetelmä käyttää eri etäisyysmittareita ja yhdistämiskriteereitä, kuten Wardin WARD.D2. Wardin WARD.D2:ssa käytetään euklidista etäisyyttä pyrkien minimoimaan ryhmien sisäinen varianssi, mikä tarkoittaa, että ryhmien sisällä olevat objektit (tässä tapauksessa diagnoositluokat ja ratkaisumäärät) ovat mahdollisimman samankaltaisia keskenään. Wardin menetelmä on hyödyllinen, kun tavoitteena on muodostaa homogeenisia ryhmiä, jotka mahdollistavat selkeän ja ymmärrettävän ryhmittelyn analysoiduista sekvensseistä.

AGNES-menetelmän käyttämä WARD.D2 -ryhmittely suoritetaan iteratiivisesti yhdistäen niitä ryhmiä, joiden yhdistäminen aiheuttaa pienimmän mahdollisen kasvun sisäisessä varianssissa. Tämä tekee menetelmästä erityisen sopivan tilanteisiin, joissa halutaan löytää luonnollisia ryhmittymiä monimutkaisista ja heterogeenisistä aineistoista. (Gabadinho ym. 2011.) Wardin menetelmä on käyttökelpoinen tämän aineiston sekvenssianalyysissa, koska se mahdollistaa sekvenssien luokittelun selkeisiin ja helposti tulkittaviin ryhmiin, ja yhdistettynä OM-etäisyyksiin tarjoaa tehokkaan tavan ryhmitellä sekvenssejä.

#### 4.2.8 Sekvenssianalyysin ja ryhmittelyn (klusterointi) laadunarviointi ja vertailu

Sekvenssianalyyseja on monenlaisia ja niitä käytetään hyvin erilaisiin tutkimuksiin, aineistoihin ja tutkimuskysymyksiin. Samalla sekvenssianalyysin laadunvarmistamiseen on tutkimuskirjallisuudessa ehdotettu erilaisia toimenpiteitä, ja niistä sopivien valitseminen voi olla monimutkaista. Tässä alaluvussa taustoitetaan ensin laadunarvioinnin ja vertailun yleisiä tapoja tuoreen kirjallisuuden ja esimerkkien kautta, ja esitellään sen jälkeen työn kannalta keskeisimmät menetelmät.

Raab ja Struffolino (2023, 46–47) edellyttävät, että tutkijoiden on sosiaalitieteiden kontekstissa lähes aina tehtävä jonkinlaista laadullista järjestelyä siihen, miten tilat kirjataan alfabetiin, ja siten astuttava hieman ulos puhtaan matemaattisteoreettisesta viitekehyksestä. Tällainen datan käsittely paremman lopputuloksen saamiseksi on siis perusteltavissa, vaikka se altistaa kritiikille datan manipuloinnin suhteen. Kun tilojen laatua pohditaan, on selvää, että vaikka useilla indikaattoreilla mitattuina esimerkkisekvenssit "sinkku(2)-yhteisasuminen(2)-naimisiinmeno(2)" ja "naimisiinmeno(2)-yhteisasuminen(2)-sinkku(2)" ovat matemaattisesti yhtäläiset, niin laadullisesti ne johtavat hyvin erilaisiin elämäntilanteisiin: henkilö päätyy naimisiin ja henkilö päätyy sinkuksi. Juuri tässä yhteiskuntapoliittinen kehys auttaa asettamaan tutkimuskysymyksen kautta laadullisia hierarkioita, ja siten perustelee astumisen ulos matemaattisesta viitekehyksestä.

Tämän aineiston eri tilojen laadullinen vertailu toisiinsa ei sinällään ole mielekäästä. Työssä ei pyritä löytämään hierarkkista järjestystä työkyvyttömyyseläkeratkaisujen ensisijaiselle lääketieteelliselle syyllä esimerkiksi siten, että tuki- ja

liikuntaelinsairaudet olisivat parempia tai huonompia kuin psykiatrian sairaudet. Kuitenkin hierarkkisesta laadullisesta järjestyksestä on hyötyä silloin, kun on tarpeen erottaa kaksi tilaa toisistaan ja tutkia näiden tilojen muutosta, kuten tehdään sairauksien syiden vaihtuessa toiseksi.

Raab ja Struffolino (2023, 48) kirjoittavat, että epävarmuusindeksi on käytännöllinen työkalu tämän kaltaiseen sekvenssien järjestyksen, tilan, keston ja laadun mittaamiseen, vaikka se altistaakin tutkimuksen kritiikille. Hieman yksinkertaista epävarmuusindeksin perusidea on painottaa siirtymiä negatiivisina tai positiivisina ja laskea näiden suhteellista osuutta saaden indeksiluvun 0 ja 1 välillä. Raab ja Struffolino (2023, 49) esittelevät myös toisen vastaavan indeksin, Brzinsky-Fayn (2007) laatuindeksin, jossa suotuisat tapahtumat kirjataan ykköseksi ja epäsuotuisat nolllaksi. Tällä laatuindeksillä aineistoista saadaan esiin suotuisten tapahtumien sarjoja. Yleisenä haasteena laatuindekseissä on tunnistettu sekä täysin muuttumattomat sekvenssit että pelkästään suotuisaksi muuttuvat sekvenssit.

Vaikka nämä indeksit ovat hyödyllisiä sekvenssianalyysin laadunarvioinnissa, Raab ja Struffolino (2023) huomauttavat, että niiden soveltaminen saattaa olla haastavaa tietyissä konteksteissa kehityskaarien tutkimiseen. Näiden laatuindeksien käyttäminen tässä työssä monimutkaistaisi tarpeettomasti tutkimusasetelmaa. Siksi tässä maisterintutkielmassa hyödynnetään hieman vanhempia, mutta valideja menetelmiä sekvenssianalyysin laadunvarmistuksessa ja vertailussa. Gabadinho ym. (2011) ja Gabadinho ym. (2013) tarjoavat kattavia menetelmiä ja R-koodiesimerkkejä, jotka ovat hyvin dokumentoituja ja laajalti käytettyjä sekvenssianalyysin yhteydessä. Näiden menetelmien avulla on mahdollista varmistaa tulosten laatu ja luotettavuus.

Gabadinhon ym. (2011) mukaan sekvenssien laadunarviointi voidaan suorittaa tarkastelemalla sekvenssien pituusjakaumaa ja varmistamalla että sekvenssit ovat johdonmukaisia ja samanpituisia, sekä tarkastamalla sekvenssoinnin jälkeen aineiston puuttuvat arvot. Vertailu tehdään laskemalla sekvenssien väliset etäisyydet käyttäen esimerkiksi OM-menetelmää ja substituutiomatriisia. Hierarkkisen ryhmittelyn osalta tuloksia voidaan arvioida silhuettianalyysin avulla, joka auttaa tunnistamaan optimaaliset ryhmät. Silhuettien leveys mittaa kuinka hyvin jokainen objekti sopii omaan klusteriinsa verrattuna lähimpään toiseen klusteriin. Leveyden arvot vaihtelevat välillä -1 ja 1, missä suuremmat arvot osoittavat paremman sopivuuden.

(Rousseuw 1987). Näiden laadunarviointimenetelmien avulla on mahdollista varmistaa sekvenssianalyysini luotettavuus ja helpottaa tulosten tulkintaa.

### 4.3 Tutkimusetiikka

Tässä maisterintutkielmassa kiinnitetään huomiota tutkimuseettisiin kysymyksiin koko prosessin ajan. Työssä pyritään noudattamaan Tutkimuseettisen neuvottelukunnan (TENK 2023) antamia ohjeita ja toimintatapoja hyvästä tieteellisestä käytännöstä, jotka korostavat avoimuuden, rehellisyyden ja huolellisuuden periaatteita. Avoimuus ilmenee prosessin ja tulosten raportoinnissa, pyrittäessä mahdollisimman hyvään läpinäkyvyyteen. Rehellisyydellä tarkoitetaan tulosten esittämistä ilman vääristelyä ja aikaisempaan tutkimukseen viittaamista asianmukaisella tavalla, kun taas huolellisuus tarkoittaa kaikkien tutkimusvaiheiden tarkkaa toteuttamista ja mahdollisten virheiden avoimuutta (TENK 2023).

Tutkimuksessa käytettiin ainoastaan Kevan rekisteriaineistoja. Henkilötietoja sisältäviä rekisteriaineistoja voidaan käyttää tutkimukseen ilman erillistä tutkittavien suostumuksen pyytämistä, jos suostumuksia ei esimerkiksi tietojen suuren määrän vuoksi ole mahdollista hankkia (Henkilötietolaki 523/1999). Aineiston käsittelyssä noudatettiin hyvää tieteellistä käytäntöä. Tietosuojasta huolehdittiin siten, että henkilöiden tunnistena käytettiin tunnistenumeroita, eivätkä esimerkiksi aineistoon kuuluvien nimet tai henkilötunnukset olleet missään vaiheessa tutkijan tiedossa.

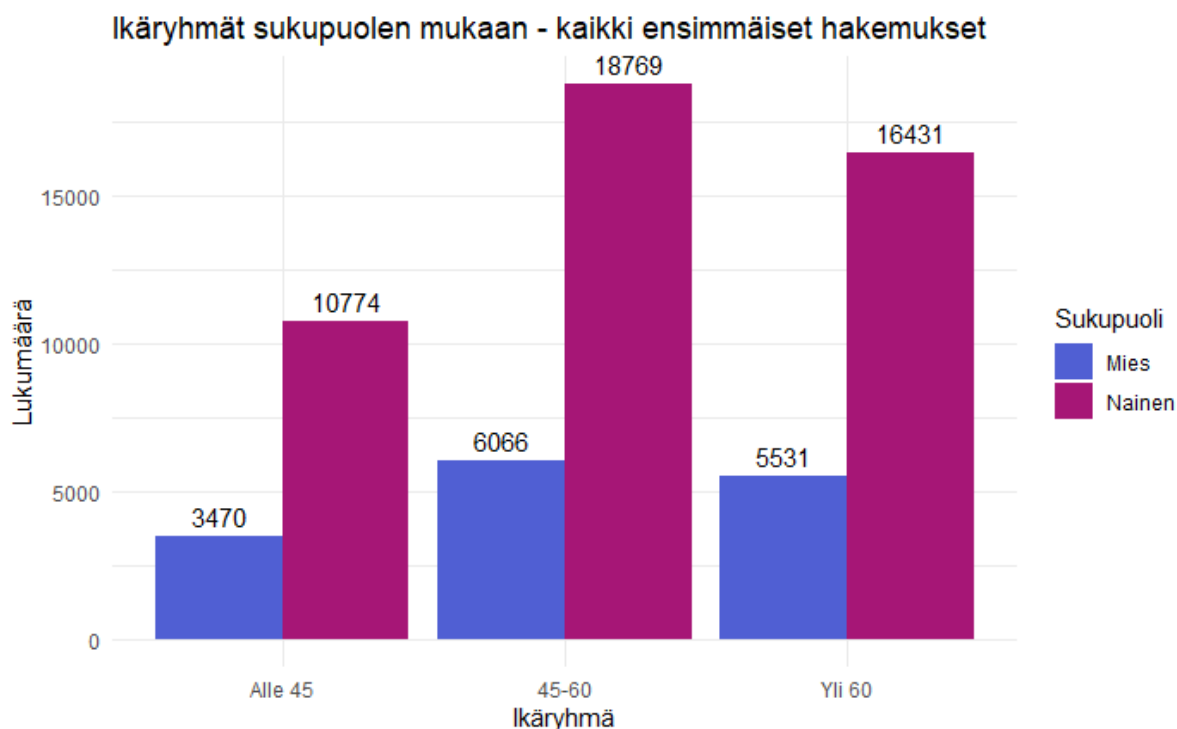
Kielimalli ChatGPT:tä (ks. liite 1, OpenAI 2024) on käytetty tässä työssä R-ohjelmistokoodin tarkistamiseen ja oikolukemiseen Helsingin yliopiston linjausten mukaisesti.

## 5 Tulokset

Tässä luvussa esitellään aineistoa ja tuloksia. Ensin käydään läpi kuvailevia tilastoja, avataan tuki- ja liikuntaelinsairauksia ja psykiatrian sairauksia tutkimusaineistossa ja käydään läpi sekvenssianalyysin tuloksia. Yleisesittelyosassa viitataan pääosin ensimmäisiin ratkaisuihin (N = 61 041), koska tämä antaa kuvan hakijajoukosta. Tekstissä, kuvioissa ja taulukoissa mainitaan erikseen tilanteet, joissa on mukana muitakin ratkaisuja. Koska valittu menetelmä on eksploratiivinen ja tarkoitus on tutkia vähemmän tunnettua aihetta, on työssä otettu tuloksiin mukaan tavallista enemmän kuviota ja taulukoita. Tämän tarkoituksena on selkeyttää niitä ilmiöitä, jotka voivat liittyä työkyvyttömyyseläkeratkaisun lääketieteellisen diagnoosin vaihtumiseen toiseksi. Taustatekijät ovat siis osa tutkielman tuloksia.

### 5.1 Aineiston yleisesittely

Tutkimusaineistossa kaikista ensimmäisistä hakijoista miehiä on neljäsosa (24,7 % eli 15 067 kappaletta) ja naisia kolmeneljäsosaa (75,3 % eli 45 974 kappaletta). Oheisessa kuviossa (Kuvio 4) näkyy ensimmäisten hakemusten ikäryhmät sukupuolen mukaan (N=61 041).



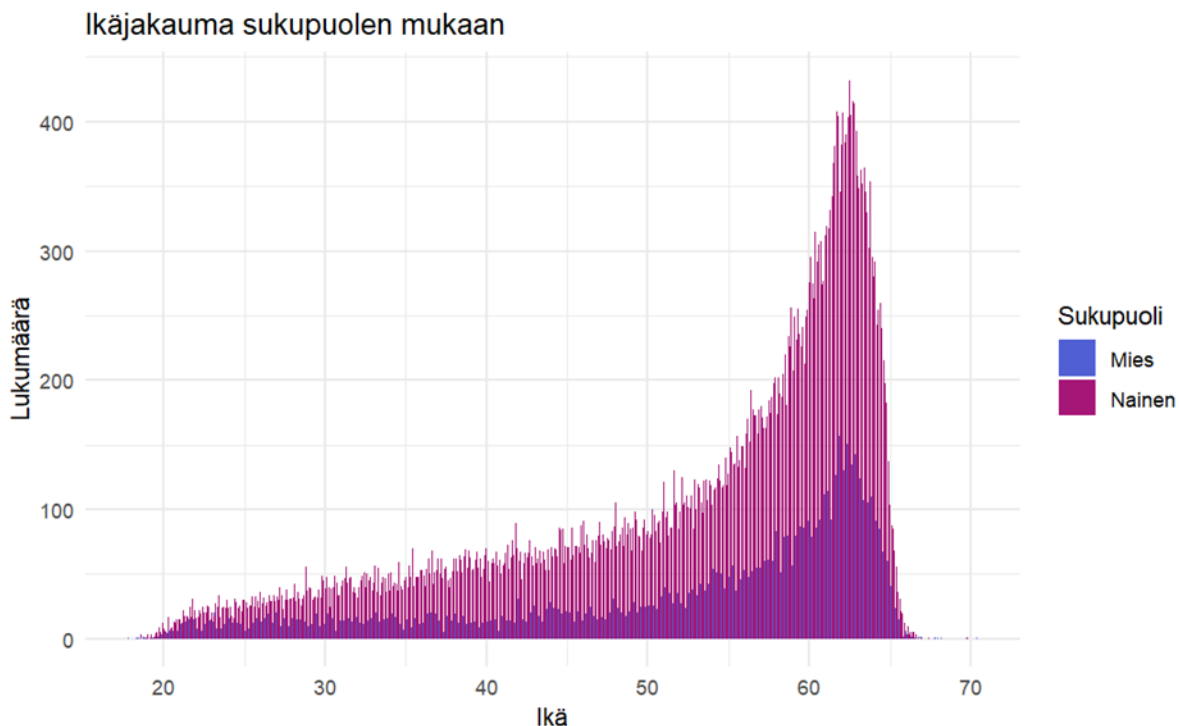
Kuvio 4. Ensimmäiset työkyvyttömyyshakemukset ikäryhmittäin ja sukupuolittain

Myös ikäryhmittäin tarkasteltuna kaikissa ryhmissä miehiä on noin neljäsosa hakijoista. Alle 45-vuotiaissa miehiä on 24,4 prosenttia, 45–60-vuotiaissa miehiä on 24,4 prosenttia ja yli 60-vuotiaissa miehiä on 25,2 prosenttia. Luvut on esitetty tarkemmin alla olevassa taulukossa (Taulukko 4).

*Taulukko 4. Ensimmäiset hakemukset ikäryhmittäin ja sukupuolittain*

<b>Ikäryhmät</b>	<b>Sukupuoli</b>	<b>Ratkaisuja (väh. 1)</b>	<b>Prosenttia</b>	<b>Yhteensä</b>
<b>Alle 45</b>	Miehiä	3 470	24,4 %	14 244
	Naisia	10 774	75,6 %	
<b>45–60</b>	Miehiä	6 066	24,4 %	24 835
	Naisia	18 769	75,6 %	
<b>Yli 60</b>	Miehiä	5 531	25,2 %	21962
	Naisia	16 431	74,8 %	
<b>Yhteensä</b>		61 041		61 041

Ensimmäisten ratkaisujen osalta koko ikäjakaumaa tarkastellessa huomataan, että miesten ja naisten ratkaisujen osuus kasvaa yhtäläisesti sekä miehillä että naisilla vanhuuseläkeikään saakka (Kuvio 5).



Kuvio 5. Ensimmäisten ratkaisujen ikäjakauma sukupuolittain

Tämä sukupuolijakauma pysyy melko samana myös seuraavien ratkaisujen kohdalla, vaikka miesten suhteellinen osuus hieman laskee ratkaisukertojen lisääntyessä.

Toisissa ratkaisuisa miehiä on 22,6 prosenttia, kolmansissa ratkaisuisa miehiä on 21,6 prosenttia ja neljänsissä ratkaisuisa miehiä on enää 21,2 prosenttia. Seuraavien ratkaisujen ikäluokkia tarkasteltaessa miesten suhteellinen osuus pysyy alle 45-vuotiaissa samana noin neljäsosana, mutta putoaa 45–60-vuotiaissa sekä yli 60-vuotiaiden ratkaisuisa noin viidennekseen ja jopa alle 20 prosenttiin kolmansista ratkaisuisa eteenpäin. Tämän aineiston perusteella voidaan sanoa, että naiset hakevat hieman useammin uusia työkyvyttömyyseläkkeitä erityisesti 45 ikävuoden jälkeen. Tämä havainto on linjassa aikaisemman tutkimuksen kanssa, joiden mukaan julkisella sektorilla naiset jäävät hieman miehiä useammin työkyvyttömyyseläkkeelle (Nivalainen ym. 2023). Havaintoon voi vaikuttaa myös naisten ja miesten erilaiset työnkuvat julkisella sektorilla.

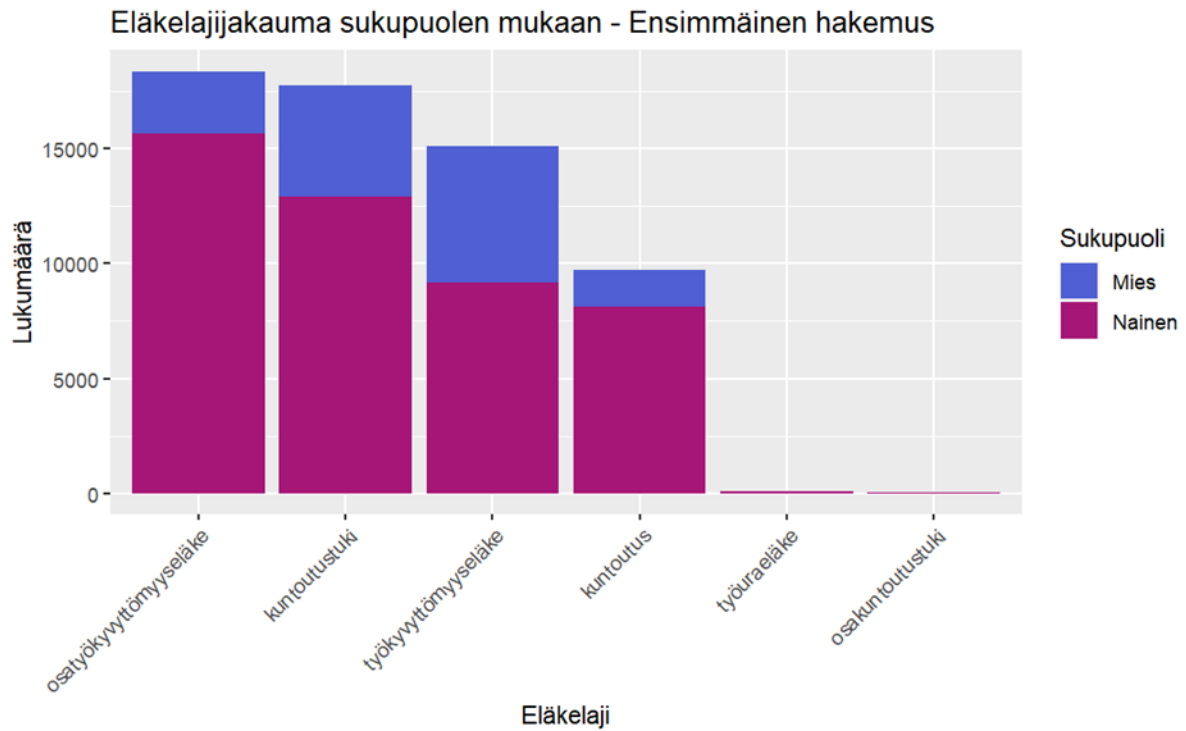
Kun tarkastellaan ensimmäisten työkyvyttömyyseläkeratkaisujen eläkelajeja sukupuolen mukaan, huomataan, että miesten ja naisten ratkaisut eroavat toisistaan. Miesten eniten hakema eläkelaji on täysi työkyvyttömyyseläke (39,3 %) kun naisilla vastaava on osatyökyvyttömyyseläke (34,1 %). Toiseksi eniten molemmissa ryhmissä haetaan kuntoutustukea (miehistä 32,1 % ja naisista 28,1 %). Tämän aineiston

perusteella havaintaan, että miehet hakevat selvästi vähemmän osatyökyvyttömyyseläkettä kuin naiset (Taulukko 5). Sukupuolijakaumia eri eläkelajien ratkaisuisissa näkyy alempana olevassa kuviossa (Kuvio 6). On hyvä huomata, että osakuntoutustukia ei juuri ilmene ensimmäisenä hakemuksena, mutta niiden hakeminen kasvaa ja tasaantuu toisesta ratkaisukerrasta eteenpäin. Toisesta ratkaisusta eteenpäin osakuntoutustukihakemuksia tulee noin 8 % kaikista ja tämä suhteellinen osuus pysyy. Myös sukupuolijakauma hakemuslajeittain pysyy likimain vastaavana kuin ensimmäisissä ratkaisuisissa.

*Taulukko 5. Hakijat sukupuolittain ja eläkelajeittain ensimmäisen ratkaisun osalta, sekä eläkelajeittain toisen ja kolmannen ratkaisun osalta.*

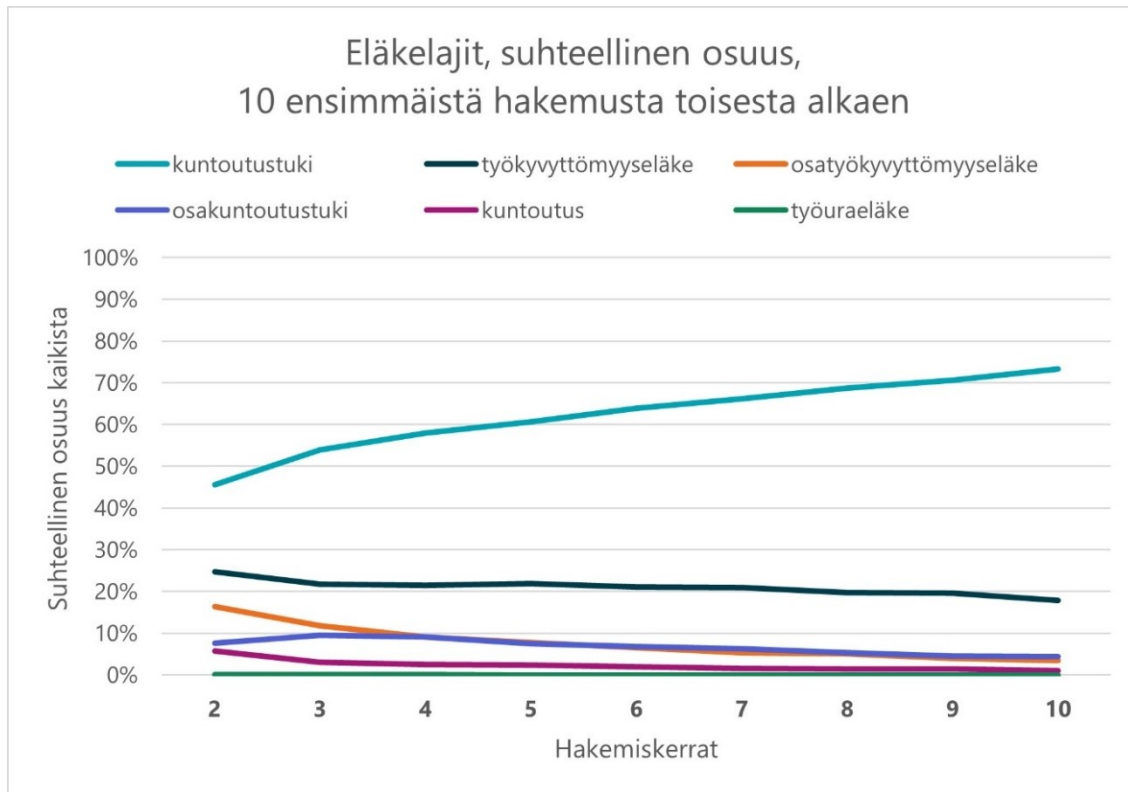
<b>Eläkelaji sukupuolen mukaan, ensimmäinen ratkaisu (N=61 041)</b>							
	<b>Osatyö- kyvyttömyys</b>	<b>Kuntoutus- tuki</b>	<b>Työkyvyttö- myys</b>	<b>Kuntoutus</b>	<b>Työura- eläke</b>	<b>Osakuntoutus- tuki</b>	<b>Yht.</b>
<b>Miehet</b>	2 703	4 829	5 921	1 577	31	6	15 067
<b>Naiset</b>	15 668	12 915	9 178	8 137	63	13	45 974
<b>Yht.</b>	18 371	17 744	15 099	9 714	94	19	61 041
<b>Miesten osuus miehistä, naisten osuus naisista (N=61 041)</b>							
<b>Miehet</b>	17,9 %	32,1 %	39,3 %	10,5 %	0,2 %	0,0 %	100,0 %
<b>Naiset</b>	34,1 %	28,1 %	20,0 %	17,7 %	0,1 %	0,0 %	100,0 %
<b>Osuus eläkelajeittain (N=61 041)</b>							
<b>Miehet</b>	14,7 %	27,2 %	39,2 %	16,2 %	33,0 %	31,6 %	
<b>Naiset</b>	85,3 %	72,8 %	60,8 %	83,8 %	67,0 %	68,4 %	
<b>Yht.</b>	100,0 %	100,0 %	100,0 %	100,0 %	100,0 %	100,0 %	
<b>Toinen ja kolmas hakemus eläkelajien mukaan (N=26 873 ja N=15 511)</b>							
<b>Toinen</b>	16 %	46 %	25 %	6 %	0 %	8 %	
<b>Kolmas</b>	12 %	54 %	22 %	3 %	0 %	10 %	

Alla olevassa kuviossa (Kuvio 6) näkyy ensimmäisen hakemuksen eläkelajit sukupuolittain.



Kuva 6. Ensimmäiset ratkaisut eläkelajien mukaan sukupuolittain

Kun tarkastellaan kymmentä ensimmäistä ratkaisua eläkelajeittain, huomataan, että kuntoutustuen osuus kasvaa, ja muiden lajien osuudet hieman vähenevät. Myös osakuntoutustuki kasvaa hieman, mutta kääntyy sitten maltilliseen laskuun. Tämä on ymmärrettävää, koska kuntoutustuet ovat määräaikaista, ja siten johtavat tyypillisemmin uusiin työkyvyttömyyseläkeratkaisuihin (Kuvio 7).



Kuva 7. Eläkelajien suhteellinen osuus, kun hakijalla on ratkaisuja enemmän kuin yksi

Kun tarkastellaan hakijoiden työkyvyttömyyseläkkeiden ensimmäisiä ratkaisuja eläkelajeittain ilman sukupuolta tai ikää, havaitaan, että suurin osa ratkaisuista on osatyökyvyttömyyseläkkeitä (30,1 %) ja kuntoutustukia (29,1 %). Tautiluokissa työkyvyttömyyden alkavuuden ensisijaisena syynä suurimpana ryhmänä ovat tuki- ja liikuntaelinsairaudet (33,9 %) ja psykiatrian sairaudet (32,6 %) (Taulukko 6).

Taulukko 6. Ensimmäiset hakemukset eläkelajeittain sekä tautiluokittain

Eläkelaji	Lukumäärä	Tautiluokka	Lukumäärä
<b>Osatyökyvyttömyyseläke</b>	18 371	<b>Tuki- ja liik. sair.</b>	20 701
<b>Kuntoutustuki</b>	17 744	<b>Psykiatrian sair.</b>	19 885
<b>Työkyvyttömyyseläke</b>	15 099	<b>Kasvaimet</b>	4 851
<b>Kuntoutus</b>	9 714	<b>Hermoston sair.</b>	4 484
<b>Työuraeläke</b>	94	<b>Verenkiertoel. sair.</b>	3 202
<b>Osakuntoutustuki</b>	19	<b>Muut sair.</b>	2 894
		<b>Vammat, myrkytykset yms.</b>	2 560
		<b>Ruuansulatusel. sair.</b>	1 472
		<b>Hengitysel. sair.</b>	992
<b>Yhteensä</b>	61 041		61 041

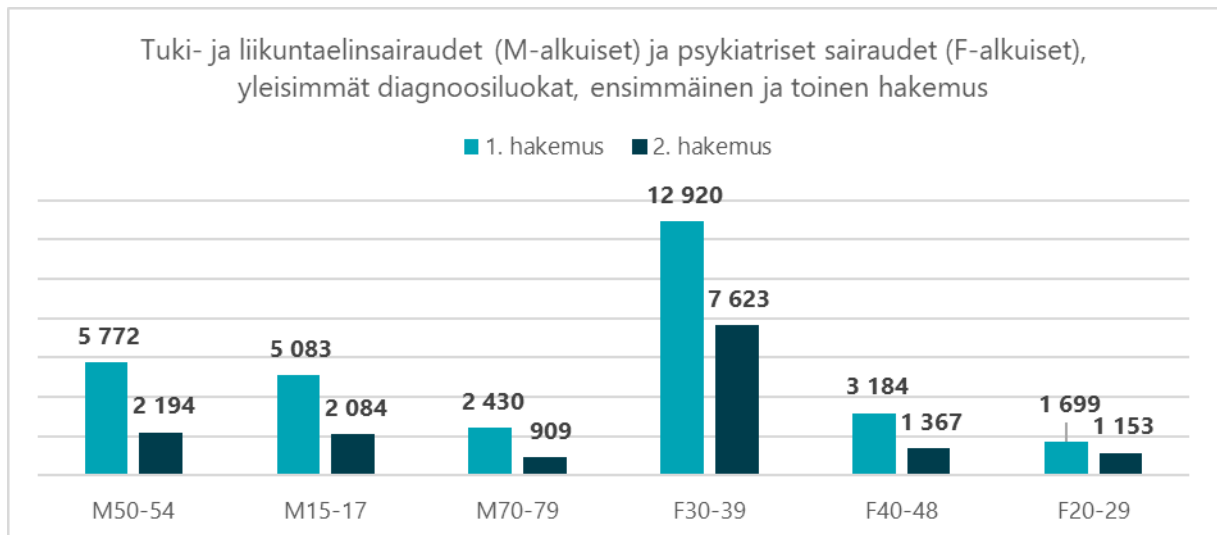
Aineistossa kaikista ensimmäisistä ratkaisuista tuki- ja liikuntaelinsairauksia on 20 701 kappaletta, joista myöntöjä oli 15 639 (75,5 %) ja hylkäyksiä 5 062 (24,5 %). Kaikkiaan tuki- ja liikuntaelinsairauksia vaihtuu psykiatrian sairauksiksi 841 kappaletta. Tämä on 4,1 prosenttia kaikista tuki- ja liikuntaelinsairauksina alkaneista. On huomionarvoista, että tuki- ja liikuntaelinsairauksia vaihtuu psykiatrian sairauksiksi selvästi enemmän kuin muiksi sairauksiksi. Esimerkiksi toisena ratkaisuna, kun päättäneet ja jatkuvat tuki- ja liikuntaelinsairausdiagnoosit on poistettu, niin psykiatrisiksi sairauksiksi vaihtuu 41,5 prosenttia kaikista vaihtuneista (Taulukko 7).

*Taulukko 7. Toinen hakemus diagnoosin mukaan, kun diagnoosi vaihtuu tuki- ja liikuntaelinsairaudesta muuksi*

Toinen diagnoosi	lkm.	%-osuus
<b>Psykiatrian sair.</b>	432	41,5 %
<b>Hermoston sair.</b>	174	16,7 %
<b>Vammat, myrkytykset yms.</b>	108	10,4 %
<b>Kasvaimet</b>	77	7,4 %
<b>Muut sair.</b>	76	7,3 %
<b>Verenkiertoel. sair.</b>	74	7,1 %
<b>Ruuan sulatusel. sair.</b>	63	6,1 %
<b>Hengitysel. sair.</b>	36	3,5 %
<b>Yhteensä</b>	1 040	100,0 %

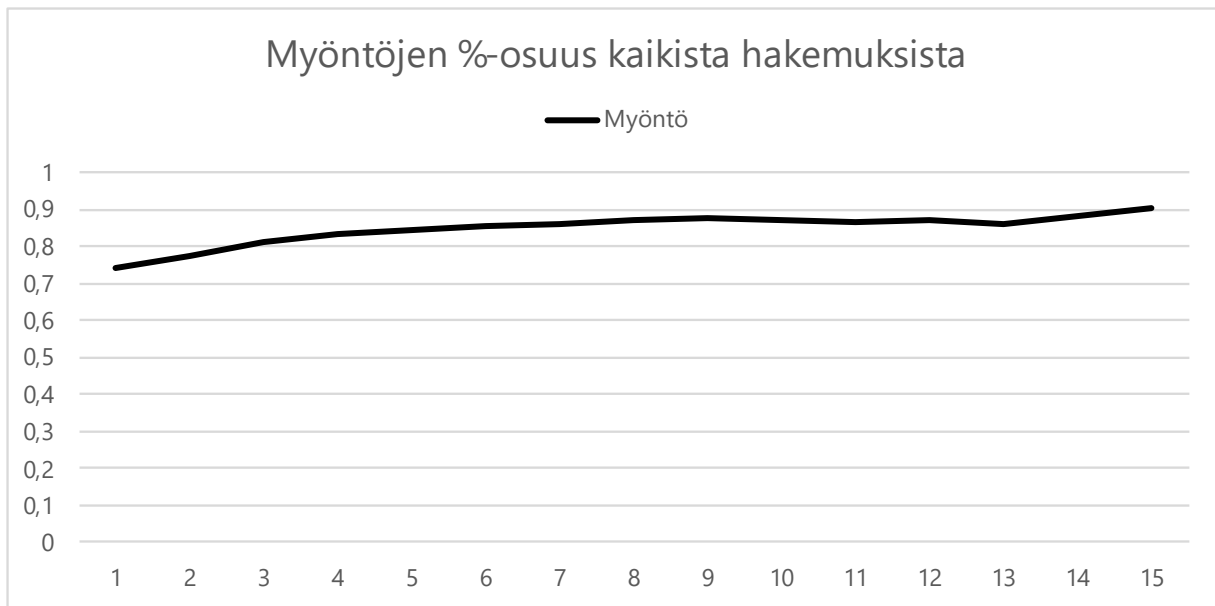
Tuki- ja liikuntaelinsairauksien alle kuuluvista diagnooseista tyypillisimpiä ovat selkään liittyvät sairaudet (M50-54; 27,9 % ensimmäisistä ratkaisuista), nivelrikot (M15-17; 24,6 % ensimmäisistä ratkaisuista) sekä pehmytkudossairaudet (M70-79; 11,7 % ensimmäisistä ratkaisuista). Ensimmäisistä ratkaisuista tuki- ja liikuntaelinsairauksien (M00-99) diagnoosien jakauma luokan sisällä on tasaisempi verrattuna psykiatrian sairauksiin (F00-99). Psykiatrian sairauksien alle kuuluvista diagnooseista useimmin esiintyvä ryhmä on mielialahäiriöt (F30-39; 65,0 % ensimmäisistä ratkaisuista), neuroottiset ja stressiin liittyvät ja somatoformiset häiriöt (F40-48; 16,0 % ensimmäisistä ratkaisuista) sekä skitsofrenia, skitsotyypinen häiriö ja harhaluuloisuushäiriöt (F20-29; 8,5 % ensimmäisistä

ratkaisuista). Psykiatrian sairauksista ensimmäisillä hakemuksilla diagnooseina näkyvät erityisesti masennustila ja toistuva masennus (F32 ja F33; 18,7 % ratkaisuista) sekä muut ahdistuneisuushäiriöt (F41; 3,2 % ensimmäisistä ratkaisuista). Kuviossa 8 näkyvät ensimmäisten ja toisten ratkaisujen tyypillisimmät diagnoosit. Kuviossa erottuu selkeästi psykiatristen sairauksien volyymit F30-39 kohdalla (Kuvio 8).



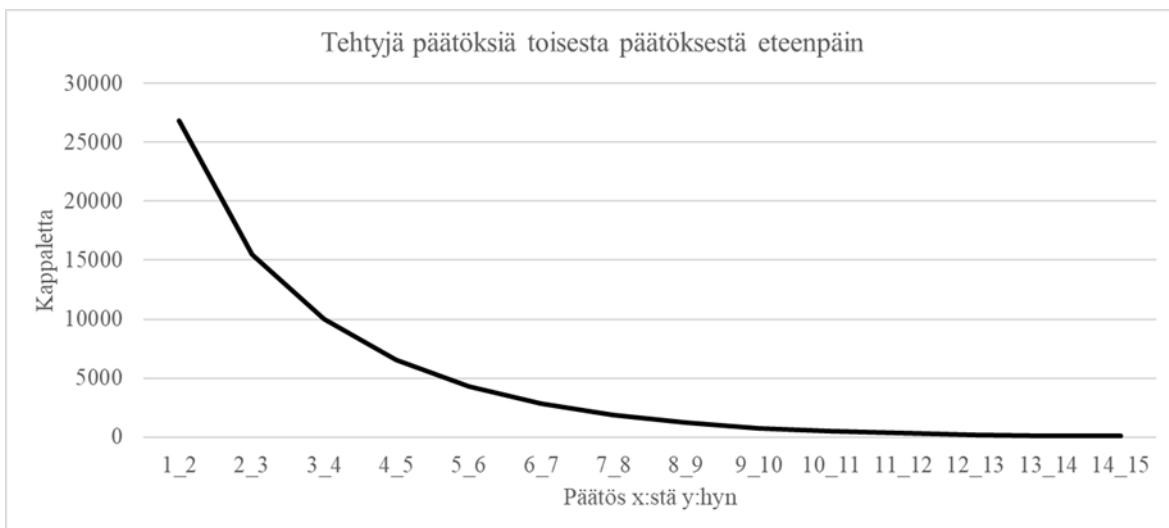
Kuva 8. Ensimmäinen ja toinen ratkaisu yleisimmissä tuki- ja liikuntaelinsairauksissa sekä psykiatrian sairauksissa

Kaikkiaan myönteisiä ratkaisuja lähtöaineistossa on 78,1 prosenttia, ja hylkääviä ratkaisuja on 21,9 prosenttia (ratkaisuja yhteensä 132 054 kappaletta). Ensimmäisistä hakemuksista myöntäviä ratkaisuja oli 74,3 prosenttia, ja seuraavista hakemuksista alkaen myöntöjen osuus kasvaa tasaisesti aina 90,4 prosenttiin asti. Tämä on ymmärrettävää, sillä tässä aineistossa seuraavassa ratkaisussa on käytännössä alla edeltävä hyväksytty työkyvyttömyyseläke. Myöntöjen osuus näkyy tarkemmin kuviossa 9 (Kuvio 9). Aineiston ja tutkimuskysymyksen kannalta ei ole tarpeen erotella myöntäviä ratkaisuja erikseen, sillä työssä tarkastellaan ensisijaisen diagnoosin vaihtumista riippumatta ratkaisusta.



Kuvio 9. Myöntöjen suhteellinen osuus ratkaisuksista per hakemuskerta

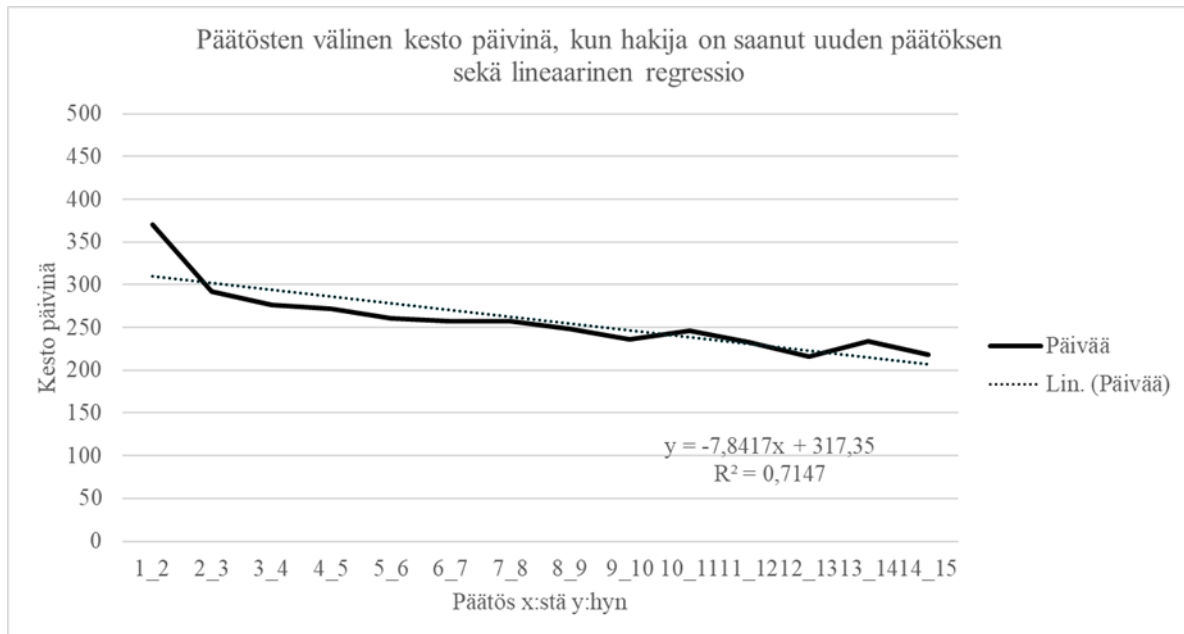
Itse hakemiskertojen määrä laskee yli ajan. Toisen ratkaisun saaneita oli aineistossa 26 816 kappaletta ja kolmannen ratkaisun saaneita enää 15 511 kappaletta. Uusien ratkaisujen määrä samalla hakijalle laskee siis melko nopeasti, ja on kuudennen ratkaisun kohdalla jo alle 3000 kappaletta. Viisitoista ratkaisua tai enemmän saaneita oli aineistossa 73 kappaletta (Kuvio 10).



Kuvio 10. Ratkaisujen määrä hakukerrasta seuraavaan (1\_2 tarkoittaa ensimmäisestä toiseen)

Ratkaisujen (sekvenssien) keston keskiarvo päivinä oli noin 260 päivää. Tämä väli on laskettu edeltävän ratkaisun alkamisajasta seuraavan ratkaisun alkamisaikaan. Keskimäärin pisin väli oli ensimmäisen ja toisen ratkaisun välillä (ka. 370 päivää).

Tämän jälkeen ratkaisun keskimääräinen väli lyheni noin kahdeksan päivää per ratkaisu kohden pysyen kuitenkin keskiarvoltaan yli 200 päivän (Kuvio 11).

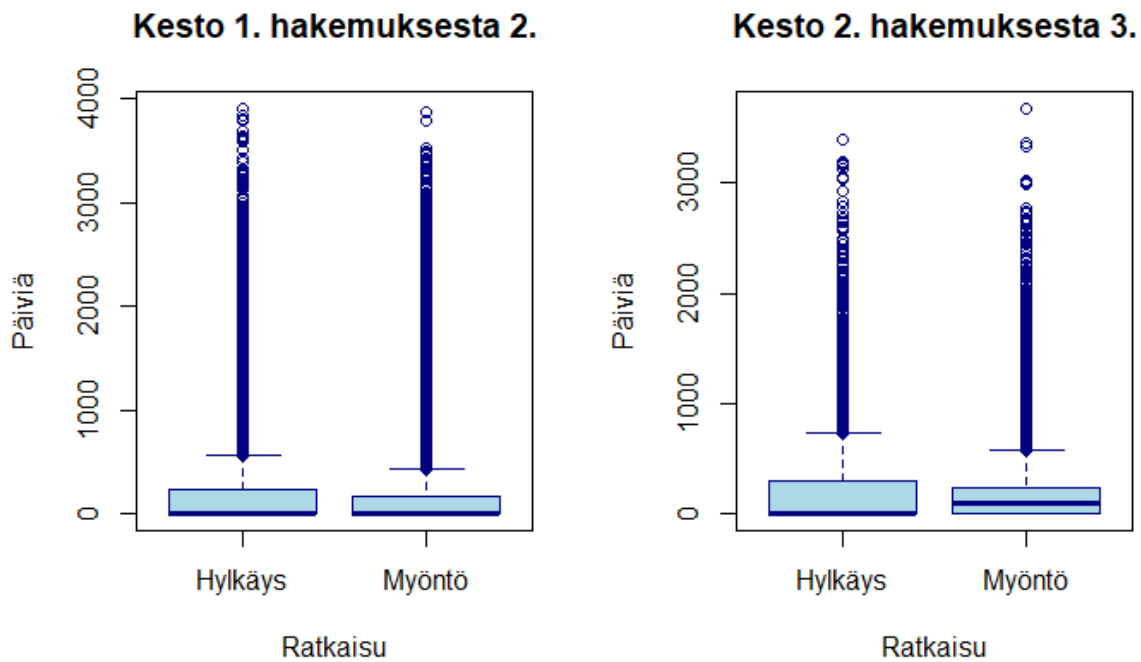


Kuvio 11. Ratkaisujen välinen keskimääräinen kesto ja lineaarinen regressio hakukerrasta seuraavaan (1\_2 tarkoittaa ensimmäisestä toiseen)

Aineistossa on kestoiltaan myös hyvin lyhyitä ratkaisuja (alle 10 päivää).

Ensimmäisen ja toisen sekvenssin välisiä alle 10 päivän tapauksia on 179 kappaletta. Kokonaisuudessa nämä ovat kuitenkin merkityksen määrää. Vastaavasti joukossa on myös erittäin pitkiä kestoja, esimerkiksi tapaus, jossa henkilö on ensin saanut hylkäävän kuntoutusratkaisun helmikuussa 2013, ja tämän jälkeen hakenut työkyvyttömyyseläkettä marraskuussa 2023, johon saanut hylkäävän ratkaisun (kestoksi tulee yhteensä 3913 päivää). Tällaiset yksittäiset poikkeamat eivät vaikuta sekvenssianalyysiin, mutta ne on hyvä huomioida, mikäli rekisteriaineiston kestoja halutaan tarkastella tarkemmin myöhemmässä tutkimuksessa.

Keskimäärin ratkaisujen välit ovat linjassa työkyvyttömyyseläkeratkaisujen yleisen luonteen kanssa. Hylkäystapausten välit ovat keskimäärin kestoiltaan hieman pidempiä kuin myöntötapausten välit lähes kaikissa ratkaisuvaiheissa. Ensimmäinen neljännes ja mediaanit ovat usein 0, mikä viittaa siihen, että monet kestot ovat lyhyitä. Keskiarvon ja kolmannen neljänneksen erot ovat myös melko lyhyitä, mikä näkyy alla olevista boxplot-kuviosta (Kuvio 12).



Kuvio 12. Boxplot-kuvio ensimmäisen ja toisen sekä toisen ja kolmannen ratkaisun välisestä kestosta päivinä

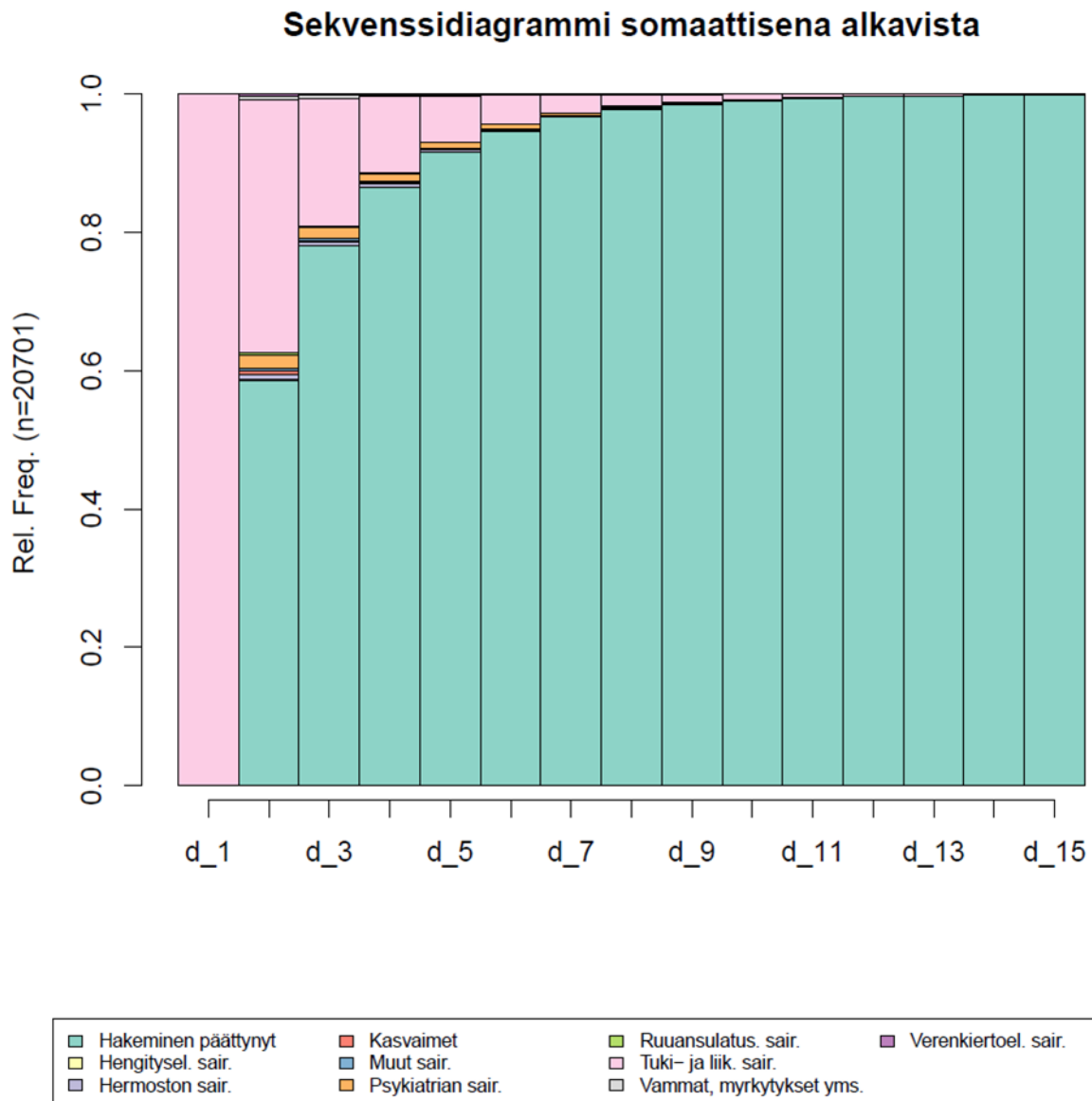
Kun tarkastellaan välien trendejä, niin huomataan että keskiarvot laskevat maltillisesti hakemukselta toiselle, mediaanit puolestaan nousevat maltillisesti. Tämä tarkoittaa sitä, että äärimmäiset arvot vähenevät, ratkaisujen väleistä (kestoista) tulee systemaattisempia ja niiden kestot pitenevät (Taulukko 8). Taulukosta näkee, että sekä hylkyissä että myönnoissä mediaani ja keskiarvo lähenevät, mutta myönnoissä molemmissa keskiluvuissa kesto kasvaa, kun taas hylkyissä keskiarvojen kestot hitaasti laskevat mutta mediaanit nousevat. Yleisellä tasolla tämä on ymmärrettävää, sillä voidaan olettaa, että työkyvyttömyyden pitkittyessä uudesta tilanteesta tulee vakiintuneempaa. Esimerkiksi pitkittyvä masennus ei luultavasti eroa kolmannen tai neljännen ratkaisun osalta kovin suuresti.

Taulukko 8. Ratkaisujen kestojen tunnusluvut päivinä hakemuskerrasta toiseen, 6 jatkohakukertaa

Hakukerta	Ratkaisu	Min.	Q1	Med.	Ka.	Q3	Max.
1_2	Hylky	0	0	0	210	228	3 913
1_2	Myöntö	0	0	0	147	170	3 879
2_3	Hylky	0	0	0	220	293	3 403
2_3	Myöntö	0	0	90	158	232	3 681
3_4	Hylky	0	0	0	211	316	2 541
3_4	Myöntö	0	0	122	174	273	3 256
4_5	Hylky	0	0	31	217	322	2 789
4_5	Myöntö	0	0	127	175	284	2 570
5_6	Hylky	0	0	53	196	307	2 594
5_6	Myöntö	0	0	124	168	274	3 005
6_7	Hylky	0	0	0	199	285	2 788
6_7	Myöntö	0	0	125	165	270	3 002

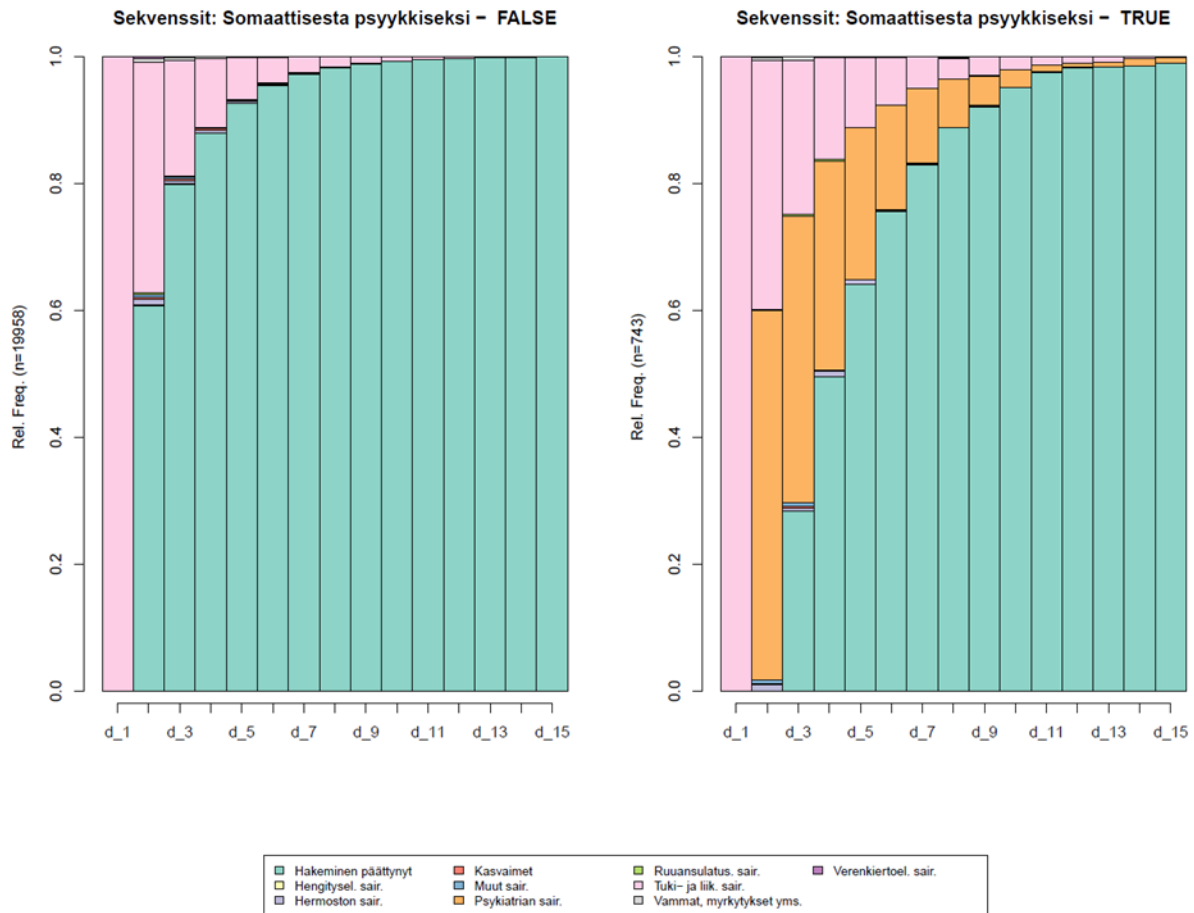
## 5.2 Sekvenssianalyysi hakemuspoluista

Koko aineistossa on kaikkiaan 132 054 kappaletta työkyvyttömyyseläkeratkaisuja. Näistä myöntävän ratkaisun saaneita on 103 100 kappaletta. Tässä työssä tarkastellaan ensimmäisistä myönnettyistä ratkaisuista (N = 61 041) niitä, jotka ovat alkaneet tuki- ja liikuntaelinsairauksina (N = 20 701). Kuviossa 14 on tilajakaumakuviona kaikki tuki- ja liikuntaelinsairauksina alkaneet ratkaisut sekä niiden jatkoratkaisut aina siihen asti, kunnes hakeminen on päättynyt (Kuvio 13).



Kuva 13. Kaikki tuki- ja liikuntaelinsairautena alkaneet hakemukset ja niiden vaihtuvuus

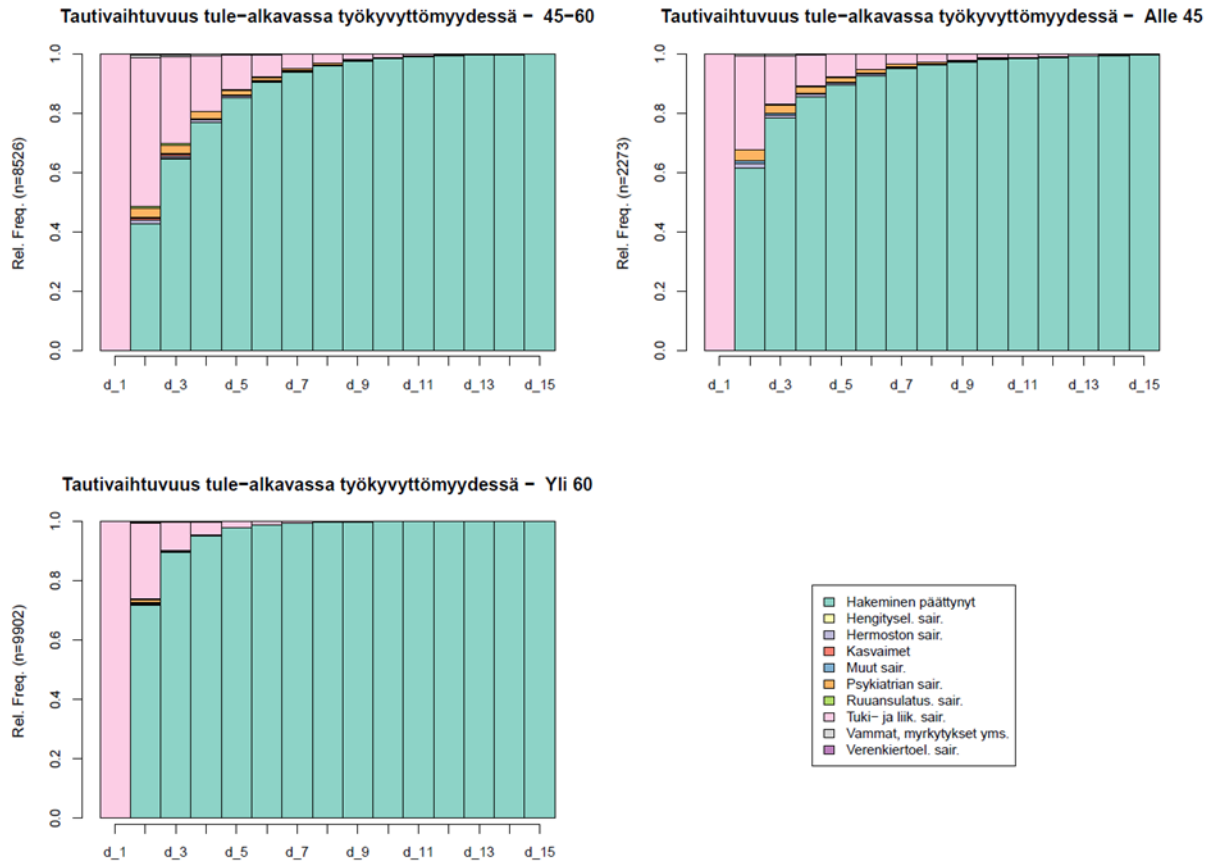
Tutkimuskysymyksen mukaisesti työssä tarkastellaan psykiatrian sairauksiksi vaihtuneita ratkaisuja. Alla olevassa kuviossa (Kuvio 14) on eritelty ne tuki- ja liikuntaelinsairautena alkaneet sekvenssit niistä, jotka vaihtuvat psykiatrian sairauksiksi (somaattisesta psyykkiseksi). Vaaleanpunaiset palkit ovat tuki- ja liikuntaelinsairauksia ja oranssit palkit ovat psykiatrian sairauksia.



Kuva 14. Tuki- ja liikuntaelinsairautena alkaneet ja psykiatrian sairauksiksi vaihtuneet hakemukset

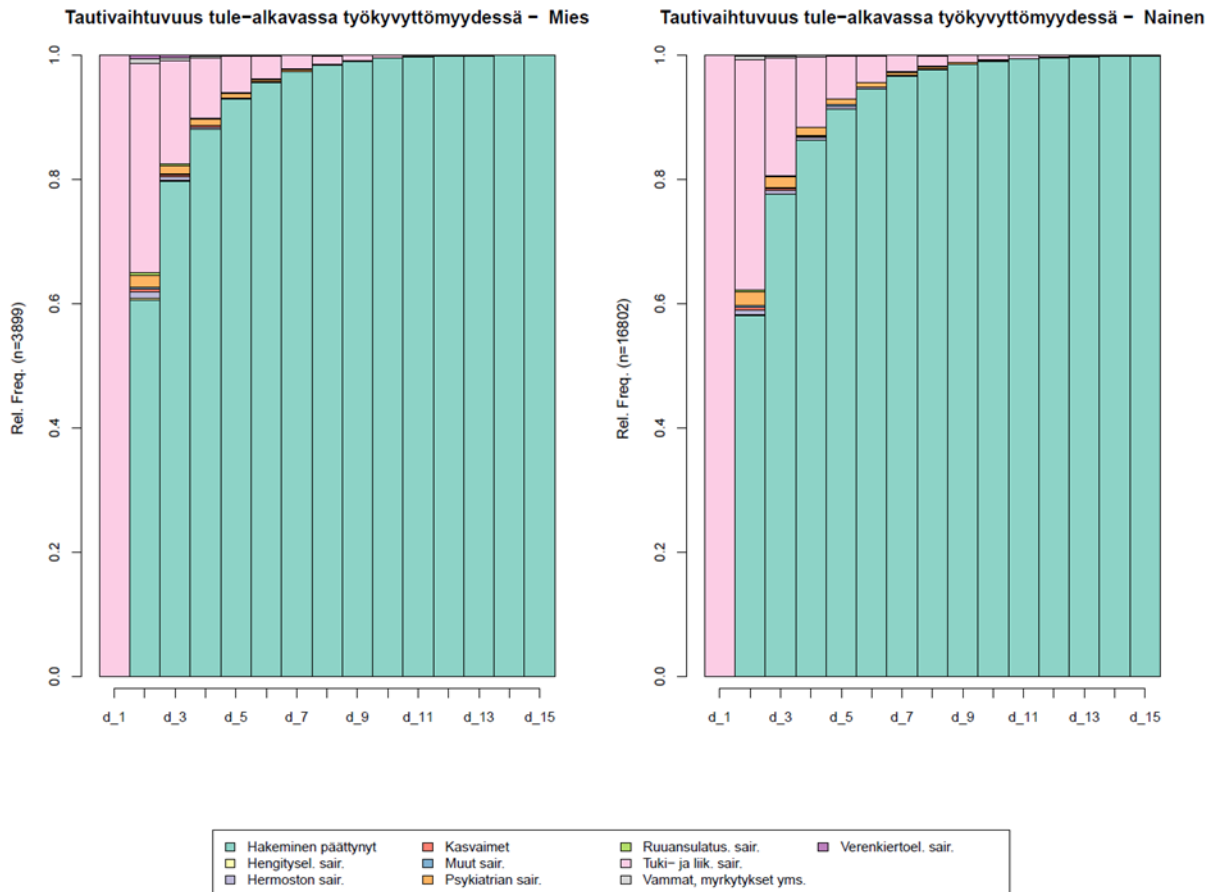
Kuviosta nähdään, että on olemassa selkeä osuus psykiatrian sairauksiksi vaihtuvia ratkaisuja. Psykiatrian sairaudet ovat vaihtuvista diagnooseista isoin erillinen joukko.

Kun sekvenssejä tarkastellaan ikäryhmittäin, niin huomataan, että erityisesti alle 45-vuotiaissa ja sekä 45–60-vuotiaissa esiintyy diagnoosin vaihtumista psykiatrian sairaudeksi. Yli 60-vuotiaissa vaihtuvuutta näkyy hieman, mutta yli 60-vuotiaissa myös jatkohakemusten määrä on vähäisempi, sillä vanhuuseläkkeen raja tulee vastaan (Kuvio 15).



Kuva 15. Tuki- ja liikuntaelinsairautena alkaneet ratkaisut eri ikäryhmissä

Sukupuolten välillä ole nähtävissä samankaltaista eroa kuin ikäluokissa. Sekvenssianalyysin hakemuspolun perusteella näyttää, ettei miesten ja naisten tautivaihtuvuuksissa psykiatrian sairauksiksi ole eroa silloin, kun ensimmäinen työkyvyttömyyseläke on ollut tuki- ja liikuntaelinsairausperusteinen (Kuvio 16).



Kuva 16. Tautivaihtuvuus tuki- ja liikuntaelinsairauksina alkaneissa hakemuksissa sukupuolittain

Aineistoa on kuitenkin tarkasteltava pienempinä otoksina ja klusteroinnin kautta, jotta todellisia eroja on mahdollista havaita.

### 5.3 Hakemuspolkujen ryhmittely (klusterointi)

Aineistosta on muodostettu erilaisia ryhmiä kahdesta ryhmästä alkaen käyttäen agglomeraattista klusterointia (AGNES). Aineistosta on mahdollista tehdä useita erilaisia ryhmittelyjä, mutta tämän tyyppisessä tutkimuksessa on kuitenkin mielekästä löytää sopiva tasapaino yksityiskohtaisuuden ja yleisyyden välillä. Aineisto ja tutkimuskysymys huomioiden kuusi ryhmää ( $k=6$ ) on sopiva määrä. Tällöin visualisointi ei ala liiaksi hävittää informaatiota, mutta tarjoaa riittävästi eroja ryhmien välillä. Ryhmien määrää kasvattaessa kuutta suuremmaksi muut ryhmät jakautuvat useampaan osaa, mutta ryhmät 3 ja 6 pysyvät yhtenäisinä. Tämä havainto tukee käytettyä ryhmien määrää.

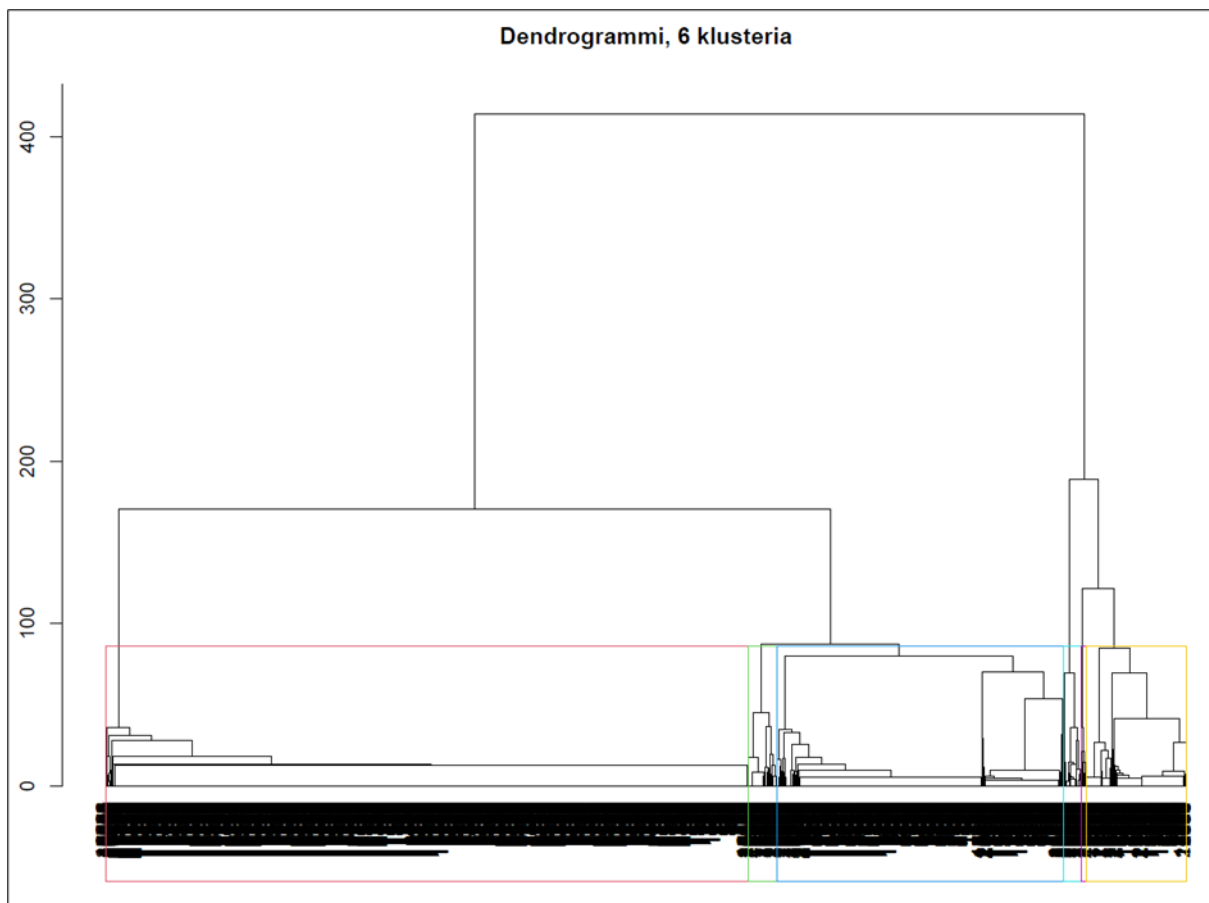
### 5.3.1 Ryhmittelyn päätulokset

Päätuloksena muodostettiin sekvenssianalyysin jälkeen kuusi ryhmää ( $k=6$ ), joiden sekvenssien jakaumat näkyvät alla olevasta taulukosta (Taulukko 9).

*Taulukko 9. Ryhmien (klusterit) sisältämät ratkaisujen lukumäärät sekä dendrogrammin ryhmän väri*

Ryhmä	1	2	3	4	5	6
Lkm.	12 310	5 489	542	342	1 912	106
Väri	Punainen	Sininen	Vihreä	Turkoosi	Keltainen	Violetti

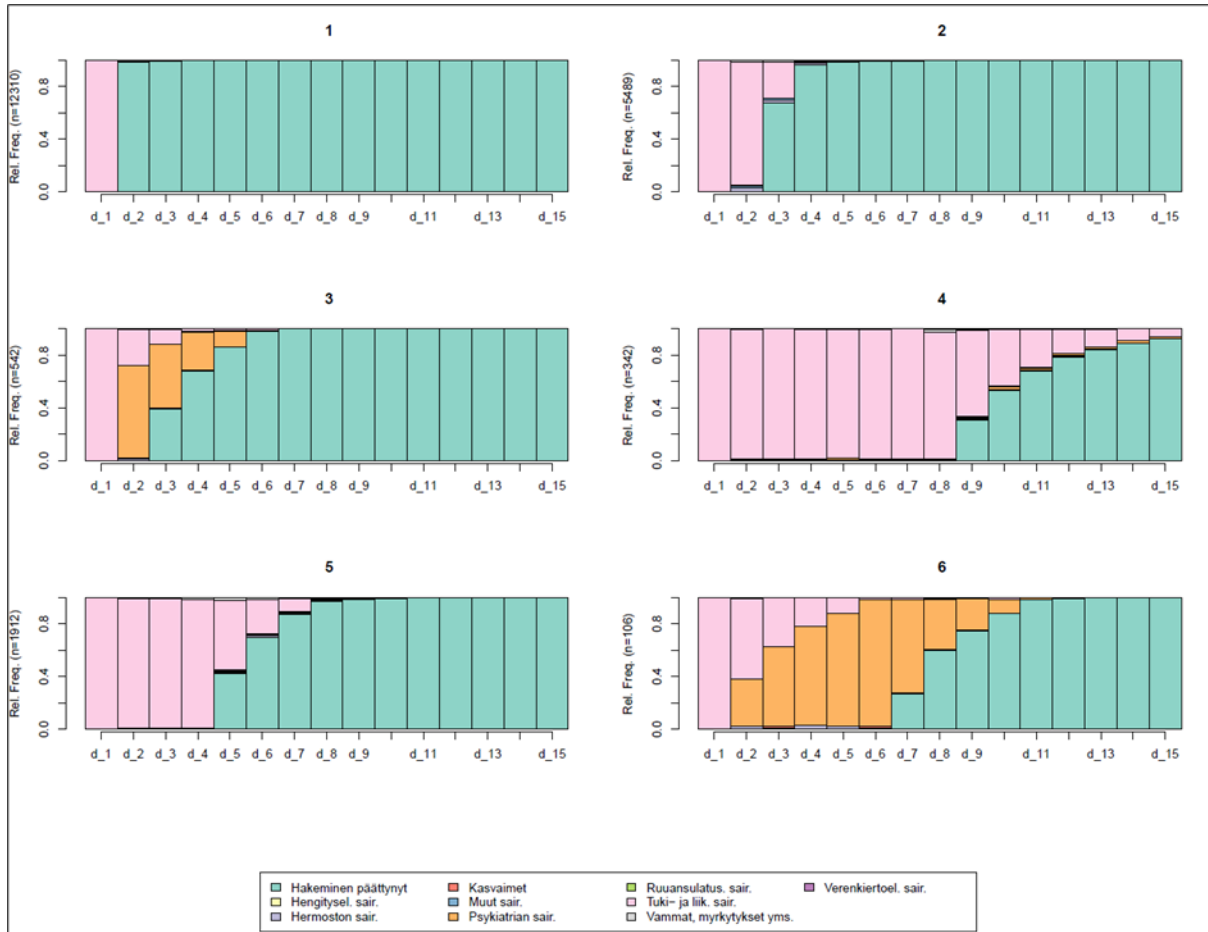
Alla olevasta dendrogrammissa nähdään ryhmien (klusterit) jakautuminen väritettyinä (Kuvio 17). Kuviossa näkyy ”bottom-up” -periaate, joka yhdistää jokaisen sekvenssin omaksi ryhmäkseen ja sitten toistuvasti yhdistää lähimmät ryhmät, kunnes kaikki sekvenssit ovat yhdessä ryhmässä.



*Kuva 17 Dendrogrammi aineiston jakautumisesta kuuteen ryhmään (klusterit) värien mukaisesti*

Wardin WARD.D2 -menetelmä osoittautui hyödylliseksi homogeenisten klustereiden muodostamisessa, ja OM-etäisyys auttoi erottamaan sekvenssien väliset muutokset.

Ryhmittelyn avulla pystyttiin tunnistamaan, missä vaiheissa ja millaisissa sekvensseissä vaihtuminen psykiatrian sairauksiksi tapahtui. Kuviossa 18 nähdään ryhmäkohtaiset hakemuspolkujen tulokset (Kuvio 18).

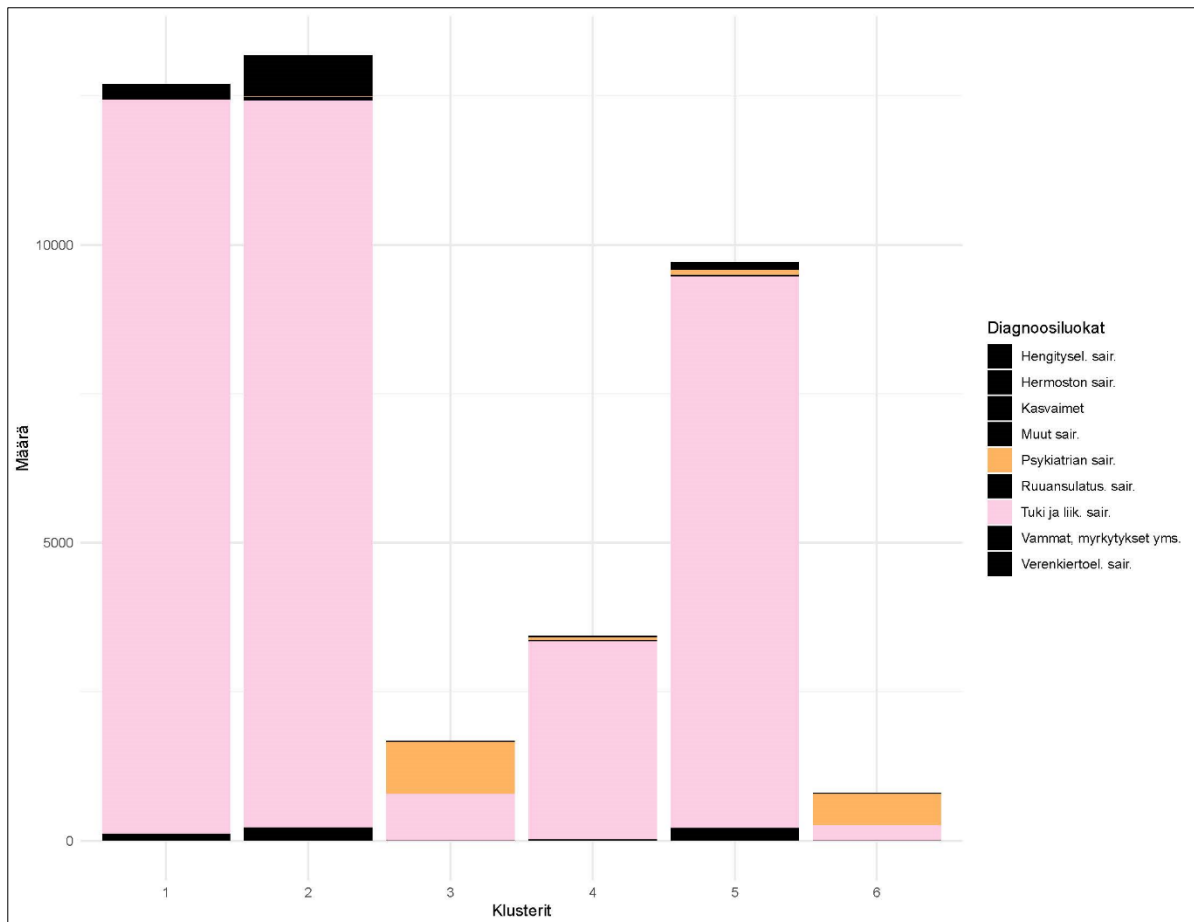


Kuva 18. Sekvenssianalyysin hakemuspolkujen tulokset ryhmittäin (klusterit)

Klusterianalyysin tuloksena on erottavissa kuusi ryhmittelyn mukaista joukkoa. Ryhmittelyssä korostuu hakemuskertojen määrä. Ryhmä 1 (N=12 310, 59 %) on suurin ryhmä, jossa yhden ratkaisun jälkeen hakeminen päättyy. Ryhmä 2 (N = 5 489, 27 %) on vähän jatkohakemuksia tehneitä mutta tuki- ja liikuntaelinsairaus ei ole merkittävässä määrin vaihtunut toiseksi. Ryhmä 3 (N = 542, 3 %) on vähän jatkohakemuksia tehneitä, mutta tuki- ja liikuntaelinsairaus vaihtuu psykiatrian sairaudeksi. Ryhmä 4 (N = 342, 2 %) on tapaukset, joissa on useita ratkaisuja, mutta ensisijainen syy ei vaihdu. Ryhmä 5 (N = 1 912, 9 %) on useita jatkohakemuksia, muttei kuitenkaan niin pitkittänyttä kuin ryhmässä 4, ja syy ei vaihdu toiseksi. Ryhmässä 6 (N = 106, 1 %) on useita jatkohakemuksia ja syy vaihtuu tuki- ja liikuntaelinsairaudesta psykiatrian sairaudeksi. Ryhmiin 4 ja 5 lukeutuu hyvin pieni osuus psykiatrian sairaudeksi vaihtuvia tapauksia.

Sekvenssianalyysin perusteella tämän aineiston hakemuksissa on erotettavissa kaksi joukkoa, joissa tuki- ja liikuntaelinsairauden ensisijainen diagnoosi vaihtuu psykiatrian sairaudeksi: lyhyemmät vaihtuvat sairastavuudet (ryhmä 3) ja pidemmät vaihtuvat sairastavuudet (ryhmä 6). Kummassakin ryhmässä suurin vaihtuvuus tapahtuu ensimmäistä ratkaisua seuraavan kolmen ratkaisun aikana.

Myös käytettäessä vähemmän ryhmiä ( $k < 6$ ) on erotettavissa samalla tavoin psykiatriset sairaudet kuin kuuden klusterin mallissa, mutta kun ryhmien määrä on 5 tai suurempi, niin erot näkyvät selkeämmin. Ryhmien määrästä riippumatta kaikissa malleissa löytyy yksi suuri ryhmä, jossa hakukerrat päättyvät ensimmäiseen ratkaisuun. Lisäksi löytyy pitkittyneiden hakemusten ryhmä, jossa on noin 300 hakemusta ja joka johtaa aina 15 ratkaisuun saakka. Kuvioista erottuu mallista riippumatta aina selvästi psykiatrian sairauksiksi vaihtuvien osuus omana ryhmänään. Kuuden ryhmän ryhmittelyllä, joka on tähän tutkimukseen sopivin määrä, voidaan tarkimmin erottaa kaksi ryhmää, ryhmä 3 ja ryhmä 6, joissa tuki- ja liikuntaelinsairaus vaihtuu psykiatrian sairaudeksi. Näiden kahden ryhmän selkein keskinäinen ero on hakemiskerroissa. Ryhmä 3:ssa hakeminen päättyy melko pian, viimeistään viidenteen ratkaisuun, kun taas ryhmä 6:ssa hakeminen jatkuu aina kymmenenteen ratkaisuun saakka. Kaikkien kuuden ryhmän työkyvyttömyyseläkeratkaisujen määrät diagnoosiluokittain näkyvät alla olevassa kuviossa, jossa on korostettuna tuki- ja liikuntaelinsairaudet (vaaleanpunainen) sekä psykiatrian sairaudet (oranssi), ja tummennettuna kaikki muut luokat selkeyden vuoksi (Kuvio 19).



Kuva 19. Ryhmittelyn (klusterit) työkyvyttömyyseläkehakemukset diagnoosiluokittain. Muut luokat tummennettu erottelun selkeydeksi.

Alla olevassa taulukossa on eroteltuna sekvenssianalyysin ryhmittelyn jälkeiset diagnoosiluokat ryhmittäin lukumäärinä (Taulukko 10). Taulukosta huomataan ryhmien erityispiirteinä ryhmien 3 ja 6 ensisijaisen tuki- ja liikuntaelinsairausdiagnoosin vaihtuminen psykiatrian sairausdiagnoosiksi. Alemmassa taulukossa on korostettuna ne joukot, joissa diagnoosien osuus on 50 prosenttia tai enemmän (Taulukko 11). Ryhmässä 3 yli puolet ja ryhmässä 6 noin kaksikolmasosaa diagnooseista vaihtuu tuki- ja liikuntaelinsairaudesta psykiatrian sairaudeksi.



### 5.3.2 Ikä ja sukupuoli ryhmissä

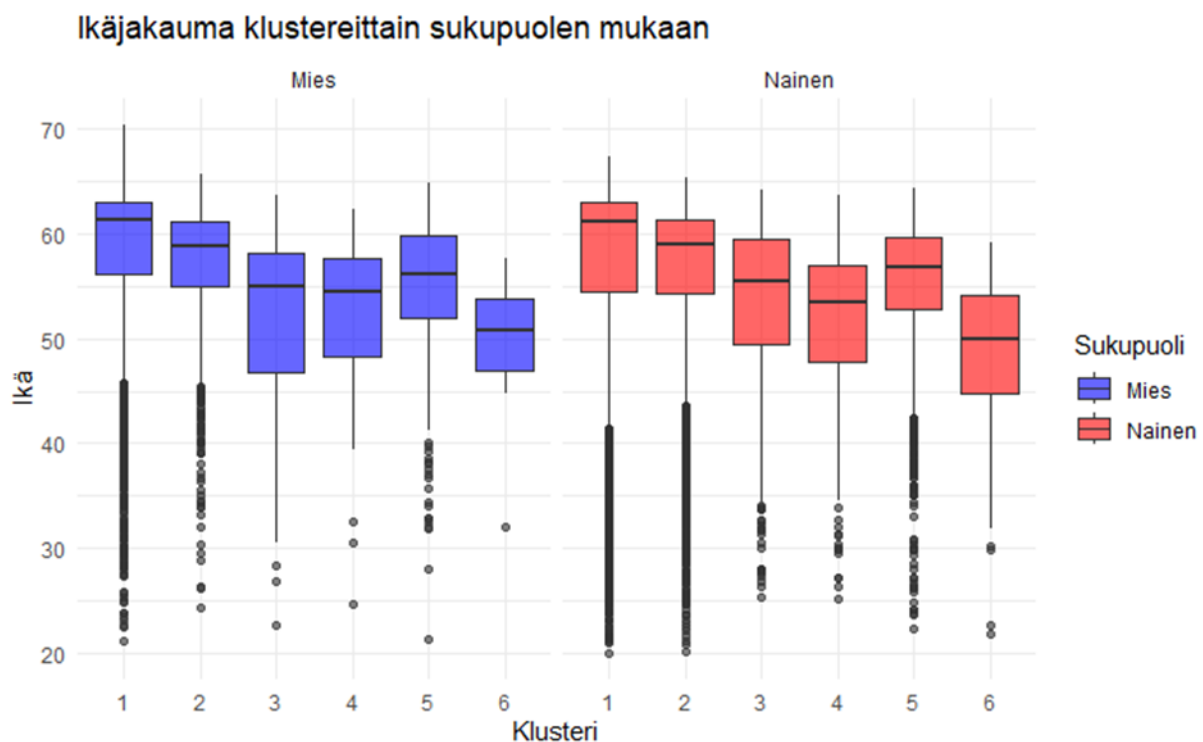
Iällä ja sukupuolella on tilastollisesti merkitsevä ero ryhmien välillä verratessa tuki- ja liikuntaelinsairautena alkaneiden hakemuksien perusteella luotuja kuutta ryhmää. Sukupuolten välinen ero ei ole sama kaikissa ryhmissä ja myös iät eroavat kaikissa ryhmissä.

Ryhmissä 1 ja 2 ikien keskiarvot ovat hieman korkeampia, ryhmässä 5 on hieman matalampi mutta silti korkeahko, ryhmissä 3 ja 4 ikien keskiarvot ovat on matalampia ja ryhmässä 6 on kaikkein matalin iän keskiarvo (Taulukko 12)

Taulukko 12. Ryhmien ikien keskiarvot

Ryhmä	1	2	3	4	5	6
Ikä ka.	57,3	56,4	52,8	51,6	55,0	48,7

Alla miesten ja naisten ikäjakaumat ryhmittäin (Kuvio 21):



Kuvio 20. Ikäjakaumat ja ryhmät eroteltuna sukupuolittain

Vaikka boxplot-kuviosta (Kuvio 20) voi olla hankala havaita eroja, niin varianssianalyysin tuloksista voidaan todeta, että ryhmät eroavat iän, sukupuolen sekä iän ja sukupuolen yhteisvaikutuksen mukaan. Ikäerot vaikuttavat tilastollisesti

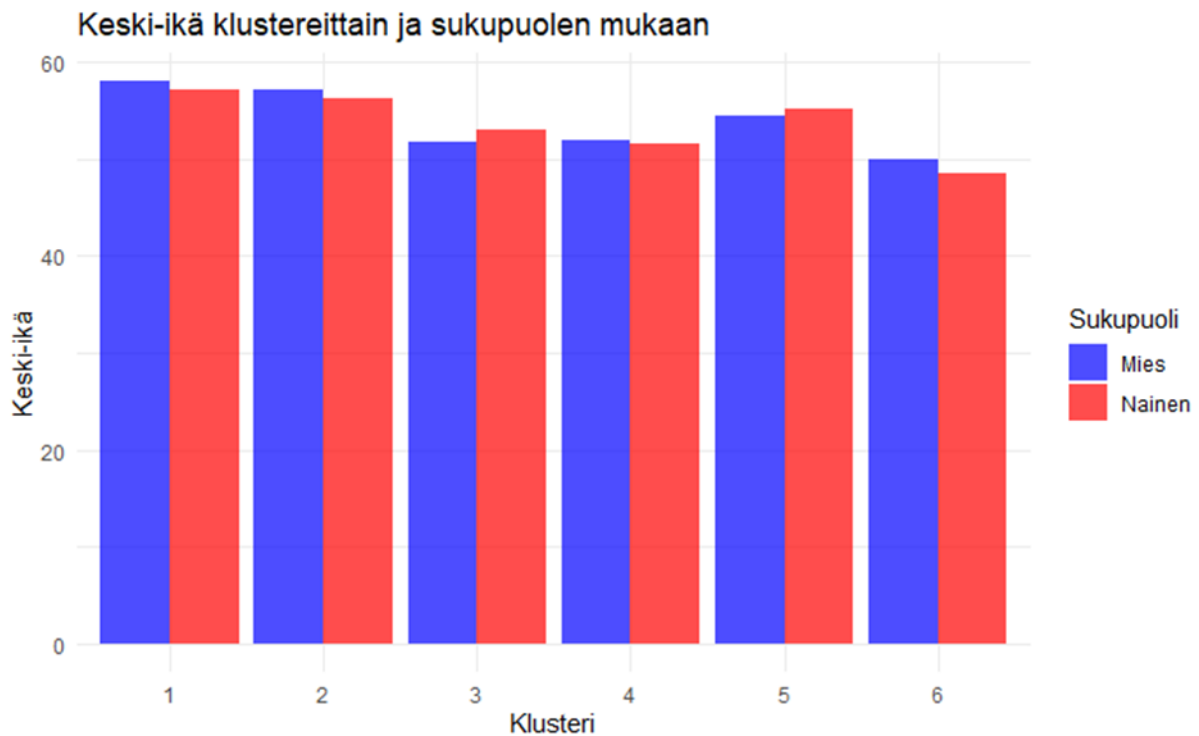
merkitsevästi ( $F(1, 20697) = 331.817, p < 0.001$ ). Tämä viittaa siihen, että ikäeroilla on huomattava vaikutus eri ryhmiin. Samoin sukupuolen vaikutus on tilastollisesti merkitsevä ( $F(1, 20697) = 20.049, p < 0.001$ ), eli sukupuolella on myös vaikutusta. Sukupuolen ja iän välinen yhteinen vaikutus on myös tilastollisesti merkitsevä ( $F(1, 20697) = 7.152, p = 0.007$ ). Eli myös ikäryhmien ja sukupuolien yhteinen yhteys vaikuttaa.

Tarkasteltaessa ryhmien ikien keskiarvoja sukupuolittain, huomataan, että myös ryhmissä keskiarvot vaihtelevat sukupuolen mukaan. Miesten ikä on keskimäärin korkeampi kuin naisten ikä ryhmissä 1, 2, 4 ja 6, mutta on matalampi kuin naisten ikä ryhmissä 3 ja 5. (Taulukko 13; Kuvio 21).

*Taulukko 13. Ikien keskiarvot ryhmittäin ja sukupuolittain*

Ryhmä	1	2	3	4	5	6
<b>Miehet</b>	58,0	57,2	51,7	51,9	54,4	49,8
<b>Naiset</b>	57,1	56,3	53,0	51,5	55,1	48,5

Tutkimuskysymyksen kannalta ryhmät 3 ja 6 osoittautuvat kiinnostaviksi. Näistä ryhmistä havaintaan, että työkyvyttömyyseläkkeiden lääketieteellinen syy vaihtuu tuki- ja liikuntaelinsairauksista psykiatrian sairauksiksi tyypillisemmin hieman varhaisemmassa iässä kuin muu vaihtuminen tai lääketieteellisen syyn jatkuminen samana. Ryhmässä 3 syy vaihtuu psykiatrian sairaudeksi miehillä keskimäärin 51-vuotiaana ja naisilla keskimäärin 53-vuotiaana sekä ryhmässä 6 miehillä keskimäärin 49-vuotiaana ja naisilla keskimäärin 48-vuotiaana. Tämä on linjassa aiemman tutkimuksen kanssa. Sekvenssianalyysin perusteella voidaan sanoa, että sekä miehillä että naisilla työkyvyttömyyden lääketieteellinen ensisijainen syy vaihtuu tuki- ja liikuntaelinsairaudesta psykiatrian sairaudeksi tyypillisemmin hieman nuoremmilla verrattuna sellaisiin tapauksiin, joissa syy ei vaihdu tai syy vaihtuu muuksi.



Kuva 21. Ikien keskiarvot ryhmittäin ja eroteltuna sukupuolen mukaan

Sukupuolittain tarkasteltuna kaikissa ryhmissä on enemmän naisia kuin miehiä (Taulukko 14).

Taulukko 14.  $\chi^2$ -testin mukaiset sukupuolen osuudet ryhmissä

Ryhmä	1	2	3	4	5	6
<b>Miehet</b>	2 413	1 001	98	48	323	16
<b>Naiset</b>	9 897	4 488	444	294	1 589	90
<b>Mies-%</b>	19,6 %	18,2 %	18,1 %	14,0 %	16,9 %	15,1 %
<b>Nainen-%</b>	80,4 %	81,8 %	81,9 %	86,0 %	83,1 %	84,9 %

Vaikka naisia on myös koko aineistossa enemmän, niin naisten suhteellinen osuus kuitenkin korostuu tuki- ja liikuntaelinsairauksista psykiatrian sairauksiin vaihtuneissa ryhmissä. Miehiä on kaikista ensimmäisistä hakemuksista 24,7 prosenttia ja vielä kuudensistakin hakemuksista miehiä on 20,5 prosenttia. Sekvenssianalyysin ryhmissä miehiä on kuitenkin vain 15,1–19,6 prosenttia, eli huomattavasti pienempi osuus.  $\chi^2$ -testi vahvistaa tätä havaintoa (Taulukko 15).

Taulukko 15.  $\chi^2$ -testin tulokset naisten suhteellinen osuus ryhmissä

Ryhmä	$\chi^2$ -arvo	p-arvo
1	136,16	< 0,001
2	103,17	< 0,001
3	10,73	0,001
4	18,94	< 0,001
5	54,29	< 0,001
6	4,7	0,03

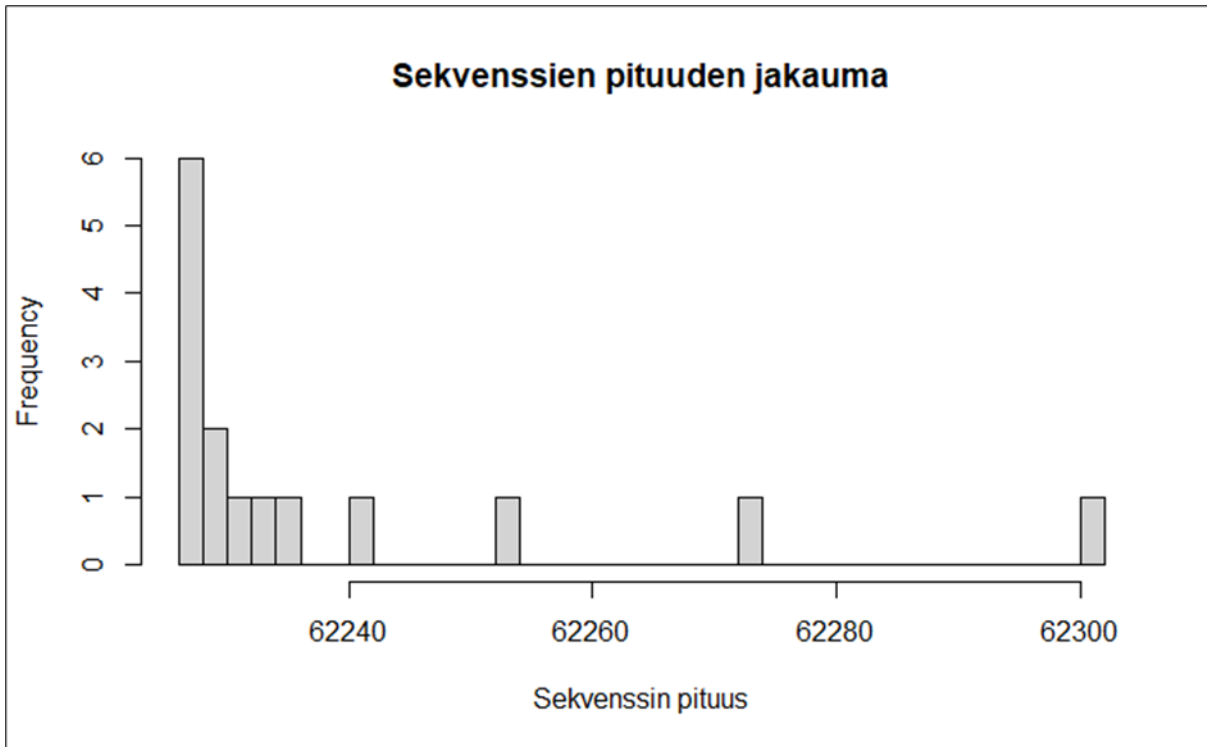
Kaikissa kuudessa ryhmässä p-arvo on huomattavasti pienempi kuin tyypillinen merkitsevyystaso (0,05), mikä tarkoittaa, että havaittujen ja odotettujen arvojen välinen ero on tilastollisesti merkitsevä. Voimme todeta, että sekvenssianalyysin ryhmien miesten ja naisten määrät poikkeavat odotetusta jakaumasta (21,4 % miehille ja 78,6 % naisille) kaikissa ryhmissä. Tuloksista huomataan myös, että ryhmät 3 ja 6 eroavat hieman muista ryhmistä. Näissä ryhmissä p-arvo poikkeaa muista jääden kuitenkin tyypillisen merkitsevyystason alle (0,05). Tähän vaikuttaa myös näiden ryhmien pieni havaintomäärä (N=542 ja N=106), joka lisää epävarmuutta tuloksissa. Tämän aineiston perusteella voidaan kuitenkin sanoa, että julkisella sektorilla työkyvyttömyyseläkkeitä hakeneista henkilöistä naisilla tapahtuu enemmän muuntumista tuki- ja liikuntaelinsairauksista psykiatrian sairauksiksi ensisijaisen lääketieteellisen diagnoosin osalta.

#### 5.4 Laadunarviointi ja vertailu

Kaikki sekvenssit ovat saman pituisia (15 tilaa). Tämä on hyvä asia sekvenssianalyysin, etäisyyslaskennan ja klusterianalyysin kannalta. Yhtenäiset sekvenssipituudet vähentävät monimutkaisuutta ja parantavat luotettavuutta. Sekvenssien väliset etäisyydet on laskettu onnistuneesti OM-menetelmällä käyttäen substituutiomatriisia TRATE-menetelmällä. Sekvenssianalyysin jälkeen puuttuvia arvoja oli 0. Sekvenssien pituudet näkyvät taulukossa sekä kuviossa (Taulukko 16; Kuvio 22).

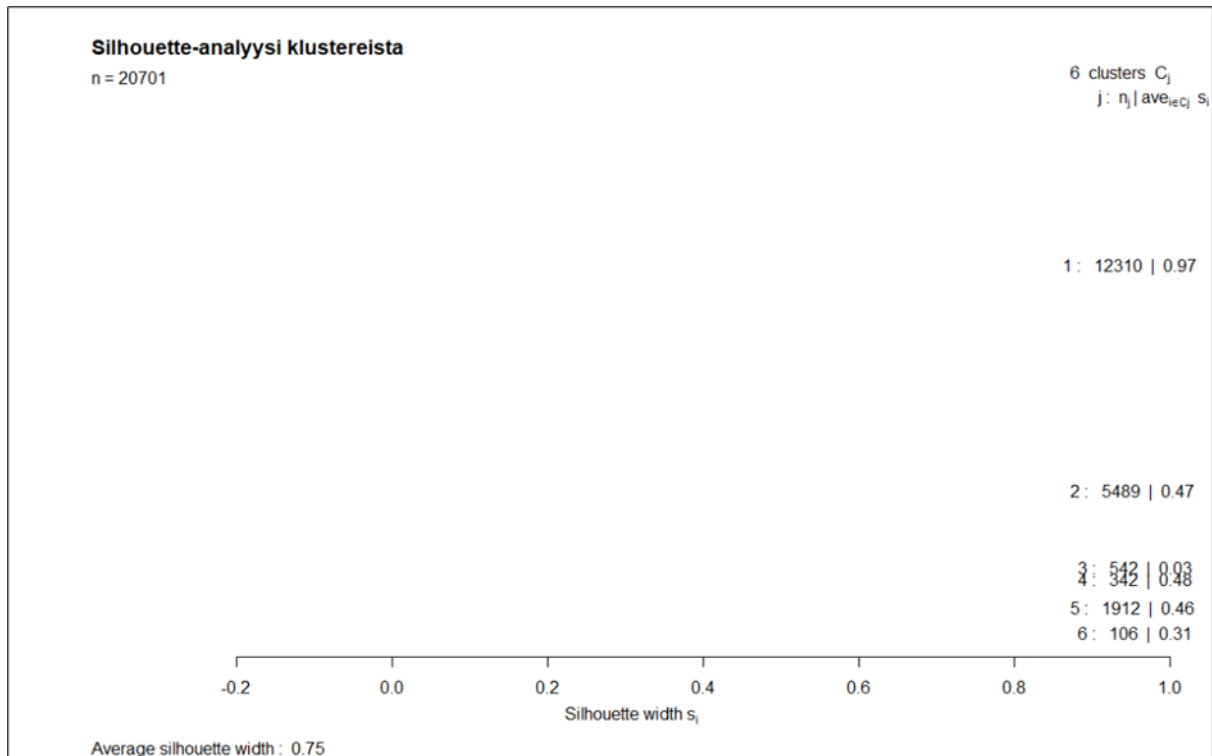
Taulukko 16. Sekvenssianalyysin laadunarvioinnissa saadut sekvenssien pituudet

Min	Q1	Mediaani	Keskiarvo	Q3	Max
62 227	62 227	62 229	62 239	62 239	62 301



Kuva 22. Sekvenssien pituuksien jakaumat

Silhuettianalyysin keskimääräinen silhuettileveys on hyvin korkea (asteikolla -1:stä 1:een tulos oli 0,75), mikä tarkoittaa sitä, että ryhmittely on onnistunut ja laatu on hyvä. Ryhmät erottuvat ja hakemuspolut ovat tulkittu mielekkäästi. Ryhmä 1 on laadultaan vahvin (0,97), ja se on muita ryhmiä suurempi. Muiden ryhmien arvot ovat hyvällä tasolla (0,31–0,47), mutta ryhmä 3:n silhuettileveys on vain 0,03, eli on melko matala. Jatkotutkimuksessa analyysia voisi esimerkiksi toistaa ilman ensimmäistä klusteria. Silhuettianalyysin tulokset näkyvät alla olevasta kuviosta (Kuvio 23).



*Kuva 23. Silhuettianalyysin tulokset*

Sekvenssianalyysin laadunvarmistuksen ja vertailun tulokset ovat hyvät ja tukevat valittua menetelmää työkyvyttömyyseläkeratkaisujen ensisijaisen lääketieteellisen syyn vaihtumisen tutkimiselle. Laadun ja vertailun näkökulmasta analyysi on luotettava, ja sen tulokset voivat tuottaa merkityksellistä tietoa työkyvyttömyyseläkeratkaisujen analysoinnissa.

## 6 Pohdinta

Tässä maisterintutkielmassa on tutkittu julkisen sektorin työkyvyttömyyseläkeratkaisujen vaihtumista tuki- ja liikuntaelinsairauksista psykiatrian sairauksiksi silloin, kun työkyvyttömyys pitkittyy ja henkilöllä on useita hakemuskertoja. Sellaiset pitkät työkyvyttömyysjaksot, joissa ei hakijalle ei tule uusia ratkaisuja ei käsitellä tässä työssä.

Tavoitteena oli tarkastella missä määrin vaihtumista tapahtuu miehillä ja naisilla sekä eri ikäryhmissä. Tutkielman tavoitteena ei ollut kattavasti ja yksityiskohtaisesti löytää kaikkia vaihtuvia diagnooseja eikä tunnistaa kaikkia muuttujia, vaan tuottaa uusia jatkotutkimusaiheita sekvenssianalyysimenetelmää hyödyntäen, ja tarkastella eksploratiivisella tutkimusotteella, onko menetelmää perusteltua käyttää jatkotutkimukseen. Tähän kysymykseen vastaus on selkeä: menetelmää on perusteltua hyödyntää jatkossakin, ja sillä voidaan löytää uusia työkyvyttömyyden joukkoja.

Menetelmä osoittautui työhön ja aineistoon sopivaksi, ja on linjassa aiemman tutkimuksen kanssa (esim. Brzinsky-Fay ym. 2006; Gabadinho 2009; Gabadinho 2013). Tehdyt valinnat olivat perusteltuja, vaikka niistä nousikin myös kriittisiä havaintoja. Näitä esitellään tarkemmin alaluvussa 6.2.

Saadut tulokset toivat myös vastauksia Garrouste ja Perdrix (2022) meta-analyysin kolmeen havaintoon. Ensinnäkin tutkimusta tehtiin avoimesti työkyvyttömyyden näkökulmasta Suomessa isoon rekisteriaineistoon perustuen, mikä helpottaa tulosten harmonisointia jatkossa esimerkiksi meta-analyyseissa. Toiseksi vaikka tuloksissa saatiin hieman tilastollisesti merkitsevää uutta tietoa, se ei ollut tämän maisterintutkielman päätarkoitus, vaan tarkoitus oli tehdä tutkimusta työkyvyttömyyteen ilmiönä sopivista vaikeasti havaittavista vaikutuksista, ja tuoda uusia mahdollisuuksia tutkimuskenttään. Kolmanneksi tutkimusta tehtiin yksityiskohtaisesti juuri reuna-alueilla. Nämä tulokset ovat tältä osin linjassa tutkimusasetelmassa asetettujen oletusten kanssa. Kuitenkin tämä työ on vain hyvin pieni askel, ja paljon täydentävää ja laajuudeltaan kattavampaa tutkimusta aiheesta tarvitaan.

## 6.1 Keskeiset tulokset

Yhtenä päätuloksena on havainto siitä, että on olemassa joukko henkilöitä, joilla tuki- ja liikuntaelinsairautena alkanut työkyvyttömyys vaihtuu psykiatrian sairaudeksi työkyvyttömyyden pitkittyessä. Vaikka ryhmä on pieni (noin 4 % tuki- ja liikuntaelinsairautena alkaneista), niin joukko on selkeästi muusta hakijajoukosta erottuva, ja koskee useita satoja ihmisiä. Erityisesti vaihtumista näkyi alle 45-vuotiaiden ja 45–60-vuotiaiden ratkaisuisissa, ja hieman enemmän naisilla kuin miehillä. Myös sukupuoli ja iällä oli yhteisvaikutusta, mutta sitä ei tämän tutkielman puitteissa ollut mahdollista käsitellä tarkemmin.

Kun tarkasteltiin ratkaisujen kestoja työkyvyttömyyden pitkittyessä, tuloksista havaittiin, että ratkaisujen väleistä tulee vähän pidempiä ja kestoiltaan tasaisempia hakemuskertojen kasvaessa. Tämä viittaa siihen, että työkyvyttömyyden pitkittyessä uudesta tilanteesta tulee vakiintuneempaa. Tämä on ymmärrettävää, sillä esimerkiksi pitkittyvä masennus ei välttämättä eroa seuraavien ratkaisujen osalta kovin suuresti, ja toistuvan sairauden diagnoosin tekeminen yksinkertaistuu.

Kun tuloksia tarkastellaan ryhmittäin, huomataan, että on olemassa selkeästi erilaisia ryhmiä. Klusterianalyysin ryhmissä yksi, kaksi, neljä ja viisi ensisijainen lääketieteellinen syy vaihtu ei psykiatriseksi sairaudeksi. Ryhmän neljä erityispiirteenä on hakemisen pitkittyminen. Siinä ratkaisuja on kahdeksasta jopa viiteentoista, kun kaikissa muissa on keskimäärin vähemmän. Ryhmissä kolme ja kuusi yli puolet ratkaisuisista vaihtuu tuki- ja liikuntaelinsairaudesta psykiatrian sairauksiksi keskimäärin viimeistään neljännen ratkaisun kohdalla, eli noin kolmen vuoden sisällä ensimmäisestä ratkaisusta. Tässä ilmiössä heijastuu varmasti myös ensisijaisen lääketieteellisen diagnoosin määrittämisen haasteet, ja sairauksien yhteisvaikutus.

Ryhmässä kolme on paljon tuki- ja liikuntaelinsairauksista psykiatrian sairauksiksi vaihtuvia, mutta työkyvyttömyys ei pitkity, eli ratkaisuja on viisi tai vähemmän. Ryhmässä kuusi työkyvyttömyys vaihtuu psykiatrian sairaudeksi ja työkyvyttömyys pitkittyy, eli ratkaisuja on vähintään kuusi. Muissakin ryhmissä on tulosten perusteella hieman psykiatrian sairauksiksi vaihtuvia, mutta niiden osuus jää hyvin pieneksi (alle 2 % ryhmän ratkaisuisista).

Tulosten perusteella Suomessa on selvästi olemassa joukko julkisen sektorin työntekijöitä, joilla työkyvyttömyyden pitkittyessä työkyvyttömyyseläkeratkaisun ensisijainen lääketieteellinen diagnoosi vaihtuu tuki- ja liikuntaelinsairaudesta psykiatriseksi sairaudeksi, ja vaihtuminen tapahtuu keskimäärin kolmen vuoden sisällä ensimmäisestä työkyvyttömyyseläkeratkaisusta.

### 6.1.1 Sukupuolten erot

Aineiston kaikista hakijoista miehiä oli noin neljäsosa ja naisia kolmeneljäsosa. Sukupuolten osuus pysyi samana kaikissa kolmessa ikäryhmässä, ja jakauma pysyi melko samana myös seuraavissa ratkaisuissa, vaikka miesten osuus hieman laskikin työkyvyttömyyden pitkittyessä. Tämä jakauma on Suomen julkiselle sektorille tyypillinen (esim. Nivalainen 2023; Laaksonen 2020; Laaksonen ym. 2022).

Tulosten perusteella miesten ja naisten tavallisimmat työkyvyttömyyseläkeratkaisut ovat keskenään hieman erilaisia. Miesten tavallisin työkyvyttömyyseläkeratkaisu on täysi työkyvyttömyyseläke (39,3 %), kun naisilla vastaava on osatyökyvyttömyyseläke (34,1 %). Miesten työkyvyttömyys on tulosten perusteella pysyvämpiluontoisempaa kuin naisten, ja naisten ratkaisuissa korostui selvästi miesten ratkaisuja enemmän määräaikaista ja osa-aikaista työkyvyttömyydet. Sukupuolijakauma hakemuskategorioittain pysyi likimain vastaavana myös ensimmäistä ratkaisua seuraavissa ratkaisuissa. Tämä tulos on linjassa aiemman tutkimuksen kanssa (kts. Kouvonen ym. 2017; Mäntö ym. 2022).

Vaikka naisia on myös koko aineistossa enemmän, niin  $\chi^2$ -testin tuloksien mukaan naisten suhteellinen osuus korostuu kaikissa tuki- ja liikuntaelinsairauksista psykiatrian sairauksiin vaihtuneissa ryhmissä. Tämä näkyy erityisesti ryhmissä neljä ja kuusi, joissa naisten osuus on jopa 85 prosenttia tai enemmän. Ryhmässä neljä tautivaihtuvuutta ei esiinny paljon, mutta hakukertoja on kahdeksasta viiteentoista. Ryhmässä kuusi on useita hakukertoja, vaikka vähemmän kuin ryhmässä neljä, mutta ensisijainen lääketieteellinen syy vaihtuu psykiatrian sairaudeksi. Erot ovat tilastollisesti merkitseviä.

Tämän aineiston perusteella voidaan sanoa, että julkisella sektorilla työkyvyttömyyseläkeratkaisun saaneista henkilöistä naisilla tyypillisemmin työkyvyttömyyseläkeratkaisut ovat osa- tai määräaikaista ja ratkaisuja tulee

useampia. Naisilla myös tapahtuu miehiä enemmän vaihtumista tuki- ja liikuntaelinsairauksista psykiatrian sairauksiksi ensisijaisen lääketieteellisen diagnoosin osalta.

### 6.1.2 Ikäryhmien erot

Ikäryhmiä tarkastellessa koko aineistossa eniten hakemuksia oli tullut 45–60-vuotiailta (47 %), toiseksi eniten hakemuksia oli tullut alle 45-vuotiailta (29 %) ja kolmanneksi eniten yli 60-vuotiailta (24 %). Kuitenkin kun tarkastellaan vain ensimmäisiä hakemuksia, niin erot kapenevat. Eniten ensimmäisiä hakemuksia oli tullut silti 45–60-vuotiailta (41 %), mutta toiseksi eniten hakemuksia yli 60-vuotiailta (36 %) ja vasta kolmanneksi eniten alle 45-vuotiailta (23 %). Ero on ymmärrettävä, sillä työkyvyttömyyden pitkittyessä moni yli 60-vuotias voi jäädä vanhuuseläkkeelle. Lisäksi aiemman tutkimuksen perusteella vanhempi väestö saattaa mieluummin sinnitellä työelämässä ollessaan lähellä vanhuuseläkeikänsä kuin aloittaa työkyvyttömyyseläkeprosessin, joka itsessään kestää yleensä kauan ja voi olla kuormittava (kts. Laaksonen 2020; Kannisto 2022).

Tulosten perusteella naiset hakevat hieman useammin uusia työkyvyttömyyseläkkeitä erityisesti 45 ikävuoden jälkeen. Tämä on linjassa aikaisemman tutkimuksen kanssa, jonka mukaan julkisella sektorilla naiset jäävät hieman miehiä useammin työkyvyttömyyseläkkeelle (Nivalainen ym. 2023). Tähän voi vaikuttaa myös naisten ja miesten erilaiset työnkuvat julkisella sektorilla (kts. Laaksonen 2020; Laaksonen ym. 2022).

Tulosten perusteella nähdään, että sekä miehillä että naisilla ensisijainen lääketieteellinen syy vaihtuu tuki- ja liikuntaelinsairaudesta psykiatrian sairaudeksi tyypillisemmin hieman nuoremmilla verrattuna sellaisiin tapauksiin, joissa syy ei vaihdu tai syy vaihtuu muuksi kuin psykiatrian sairaudeksi. Tätä havaintoa tukee se, että mielenterveydelliset syyt työkyvyttömyyden taustalla ovat kasvaneet erityisesti nuoremmissa ikäryhmissä. Käytetyn aineiston perusteella ei päästä tarkastelemaan erikseen alle 45-vuotiaiden ryhmän jakaumia. Vaikka nuoria aikuisia on jo tutkittu (esim. Heinonen ym. 2022), niin tämä on kuitenkin tärkeä jatkotutkimuksen aihe, jota voisi tarkastella sekvenssianalyysin avulla.

Ikäryhmien näkökulmasta ryhmä kuusi eroaa selvimmin niistä hakijoista, joilla hakeminen päättyy ensimmäiseen tai toiseen ratkaisuun. Ryhmässä kuusi hakukertoja on useita ja työkyvyttömyyden peruste vaihtuu psykiatriseksi. Tämän ryhmän ikien keskiarvo on vain 48 vuotta, joten keskimäärin tämän ryhmän jäsenet olisivat työelämässä parhaimmillaan vielä noin 20 vuotta. Myös ryhmässä kolme ja neljä ikien keskiarvot jäävät alle 52 ikävuoden kun ryhmissä yksi, kaksi ja viisi keskiarvo on yli 55 vuotta. Ryhmissä neljä ja kuusi on useita hakukertoja, ja matalampaa ikää voi selittää hakemuspolun alkaminen nuorempana siinä missä vanhemmille hakijoille ei ehdi kertyä niin montaa hakukertaa. Tämä ei kuitenkaan päde ryhmään kolme, jossa hakukertoja on vähän ja keski-ikä jää silti alemmaksi.

Tulosten perusteella voidaan sanoa, että tautivaihtuvuudella tuki- ja liikuntaelinsairaudesta psykiatriansairaudeksi on olemassa yhteys hieman nuorempaan ikään. Myös sukupuolen ja iän yhteisvaikutuksesta oli tuloksissa viitteitä. Tätä on syytä tarkastella jatkotutkimuksessa.

## **6.2 Tutkimuksen luotettavuus ja kriittinen tarkastelu**

Maisterintutkielman aineistona käytettiin Kevan rekisteriaineistoa, jota voidaan pitää julkista sektoria edustavana. Aineistoon kuului rajausten jälkeen vuosina 2013–2023 työkyvyttömyyseläkeratkaisun saaneita henkilöitä 61 041 henkilöä, joista useilla oli monia ratkaisuja. Laajan rekisteriaineiston hyötynä on se, ettei tutkimusjoukkoon valittujen satunnaisuus ole haaste, eikä tutkimukseen liity esimerkiksi kyselytutkimuksille tyypillistä vastauskatoa. Taustamuuttajat ovat peräisin luotettavasta rekisteristä, eikä sisällä puuttuvia arvoja. Työkyvyttömyysratkaisut ovat hoitavan lääkärin, vakuutuslääkärin ja ratkaisuasiantuntijan toteamia ja perustuvat itseraportoinnin sijaan rekisteriaineistoon. Etuna voidaan pitää myös sitä, että Kevan rekisteriaineistossa on henkilöitä kansallisesti koko Suomesta, eri ikäryhmistä ja sukupuolista. Naisvaltaisia aloja tutkittaessa miesten osuus jää usein pienemmäksi, näin ei kuitenkaan tämän aineiston osalta tapahdu, vaan kaikki ikäryhmät ja molemmat sukupuolet ovat hyvin edustettuina. Aineiston etuna on myös se, että aineisto on melko helposti saatavilla, mutta työn haasteena on vähemmän tunnetut ja osin monimutkaiset sekvenssianalyysi- ja klusterianalyysimenetelmät sekä niiden tuottamat hakemuspolut.

Rekisteriaineistojen haasteina puolestaan voidaan pitää aineiston vaatimaa käsittelyä ja tarkastamista virheiden poistamiseksi ja aineiston muokkaamiseksi tutkimuskäyttöön soveltuvaksi, tietojen kattavuutta sekä muuttujien rajoittuneisuutta. Esimerkiksi aineistossa sukupuoli on jaoteltu binäärisesti miehiin ja naisiin. Tämä jaottelu tuottaa tunnistettuja haasteita sukupuolten moninaisuuden ja vähemmistöjen näkökulmasta, ja näihin ongelmiin on syytä paneutua jatkotutkimuksessa. Aineiston käsittely tutkimuskäyttöön soveltuvaksi poistaa väistämättä osan informaatiosta.

Menetelmällisesti tämän maisterintutkielman luotettavuuteen vaikuttavista tekijöistä voi mainita, että sekvenssianalyysi ja erityisesti optimal matching ovat sosiaalitieteissä verrattain vähän tunnettuja, ja niitä on myös kritisoitu (kts. Aisenbrey & Fasang 2010; Raab & Struffolino 2023, 13). Menetelmää käytettäessä tutkijan analyttiset valinnat esimerkiksi aineiston käsittelyyn, aikapisteiden ja sekvenssoinnin osalta vaikuttavat lopputulokseen. Analyysin esivalmistelun tulee olla huolellisesti tehty ennen kuin sekvenssianalyysia aletaan suorittaa. Alan parhaat käytännöt tulevatkin esiin vähitellen menetelmien tullessa tutummaksi. Tässä maisterintutkielmassa Alfabetin tilat on huolellisesti valittu, ja välien määrittely on ollut selkeää, koska jokaisessa ratkaisussa ja sitä seuraavassa ratkaisussa on aina ratkaisupäivämäärä, eikä tutkimuskysymyksen kannalta vaihtuvuutta tarkasteltaessa ole merkitystä tapahtuuko muutos kahden kuukauden vai kahden vuoden päästä.

Viime vuosina laskentateho on parantunut ja sekvenssianalyysin soveltaminen on laajentunut. Perinteisemmin menetelmää on sovellettu esimerkiksi monimutkaisen DNA:n selvittämiseksi tapauksissa, joissa voi olla miljoonia markkereita, joiden vähäinenkin ryhmittely (klusterointi) tuottaa suuria epätarkkuuksia. Sosiaalitieteen kysymyksissä aineiston laadukkaalla alustuksella päästään yksinkertaisempaan sekvenssianalyysiin, jolloin tulosten tulkinta yksinkertaistuu, selkeytyy ja taustalla olevalta matematiikaltakin voidaan odottaa vähemmän. Kun tarkastellaan työkyvyttömyyseläkeratkaisuja rekisteriaineiston perusteella, tiedetään, että ne ovat aina joko hylättyjä tai myönnettyjä, ja niissä on aina ensisijainen lääketieteellinen syy. Tällöin on yksinkertaista järjestää jonoon ryhmiä, joissa syy vaihtuu toiseksi. Tämä mahdollistaa selkeän sekvenssianalyysin ja hakemuskäytön visualisoinnin tarkasteltaessa työkyvyttömyyseläkkeiden ensisijaisia lääketieteellisiä syitä yli ajan.

Maisterintutkielman luotettavuutta voidaan arvioida aineiston, mittareiden ja menetelmien ohella myös laajemmalla tasolla suhteessa tavoitteisiin ja tarkastella tutkimusasetelmaan liittyviä valintoja. Ensinnäkin on hyvä pohtia, minkälaisen tutkimusjoukon työkyvyttömyyseläkeratkaisuja on aineiston osalta mahdollista kuvata. Tässä työssä tarkastellaan ainoastaan henkilöitä, jotka saavat ensimmäisen ratkaisun jälkeen lisää ratkaisuja. Tällöin tarkastelun ulkopuolelle jäävät esimerkiksi kaikki sellaiset henkilöt, jotka ovat jo siirtyneet pois työelämästä eikä heillä ole syytä hakea uutta ratkaisua, vaikka työkyvyttömyyden alkuperäinen lääketieteellinen syy tosiasiassa vaihtuisikin. Poisvalikoituminen voi vaikuttaa tulosten yleistettävyyteen, kun tiedetään, että eri ikäryhmissä ja sukupuolissa työkyvyttömyyden perusteet ovat muutenkin hieman erilaisia. Toiseksi ammattiluokan, koulutuksen ja sosioekonomisen aseman puuttuminen tutkimusasetelmasta jättää selkeitä aukkoja tulosten yleistettävyyteen, erityisesti kun tiedetään, että työkyvyttömyyseläkkeelle siirtyminen johtaa fyysisen toimintakyvyn heikentymiseen. Pitkäaikainen altistuminen fyysisesti raskaalle työlle lisää tuki- ja liikuntaelinsairauksien sekä ennenaikaisen kuoleman riskiä. (Mänty ym. 2022; Ervasti ym. 2018.) Kolmanneksi aineiston henkilöt ovat julkisen sektorin työntekijöitä suomalaisessa työkyvyttömyyseläkekontekstissa, ja tulosten yleistettävyyttä kansainvälisesti on pohdittava tarkasti. On myös hyvä huomioida, että Suomessa pysyvien ja määräaikaisten työkyvyttömyyseläkkeiden määriin voi vaikuttaa se, että työsuhde on helpompi irtisanoa, mikäli henkilö todetaan pysyvästi työkyvyttömäksi. Siksi työntekijällä voi olla kannustin yrittää pysyä määräaikaisella tai osa-aikaisella eläke-etuudella, vaikka pysyvän työkyvyttömyyden perusteet olisivatkin olemassa.

Työssä on tunnistettu myös holistisen lähestymistavan tarve työkyvyttömyystilanteissa, joissa monia sairauksia käsitellään samanaikaisesti. Diagnostiikan haasteet vaikuttavat työkyvyttömyyseläkkeiden myöntämiseen, sillä työkyvyttömyyseläkeratkaisuissa merkitään vain yksi ensisijainen lääketieteellinen syy. Potilaan diagnoosi voi olla ylimalkainen, potilas saattaa olla haluton tai kykenemätön kertomaan psyykkisistä haasteistaan, psykiatristen sairauksien stigma voi vaikuttaa eri ikäluokkiin eri tavoin ja annettu diagnoosi voi olla alun perinkin virheellinen, jos perimmäisiä syitä ei tunnisteta. Myös etävastaanottojen yleistymisen tuo mukanaan uusia haasteita. Tämän lisäksi diagnoosiluokkiin perustuvaan työkyvyttömyyseläkeratkaisujen sekvenssianalyysiin liittyen on tärkeä havaita, että ICD-10-luokituksessa yksittäiseen pääluokkaan luetaan etiologialtaan

hyvin erilaisia sairauksia. Siksi esimerkiksi havaituista sukupuolieroista ei voida tehdä tarkkoja johtopäätöksiä yksittäisestä sairaudesta johtuvien työkyvyttömyyseläkeratkaisujen diagnoosien sukupuolieroista.

Tässä maisterintutkielmassa käytetyn rekisteriaineiston perusteella ei myöskään päästä tarkastelemaan sitä, mitä työkyvyttömyyseläkeratkaisujen välillä tapahtuu. Tulosten perusteella ei voida sanoa miksi joku ratkaisu seuraa toista, eikä päästä tarkastelemaan miten esimerkiksi työhön paluut tai työkokeilut ovat vaikuttaneet ratkaisujen välillä. Tämän kaltaisessa työssä ei suoraan tulosten osalta päästä tarkkailemaan sitäkään, mitä yhteiskunnallisia muutoksia pitkien tarkastelujaksojen aikana on tapahtunut, vaan tärkeimmät tapahtumat on muilla keinoin saatettava tulosten kanssa yhteen.

Tarkastelutavan vuoksi myös hakemuspolkujen yksityiskohdat ja syy-seuraussuhteet jäävät täysin epävarmoiksi. Yhdessä esimerkissä henkilö voi loukata itsensä liikennetapaturmassa ja saada siksi tuki- ja liikuntaelinsairauden perusteella määräämäänsä työkyvyttömyyseläkettä, ja vuosia myöhemmin henkilö voi sairastua masennukseen. Toisessa esimerkissä henkilö voi psykiatristen ongelmien vuoksi ajautua vahingoittamaan itseään vakavasti ja saada siksi tuki- ja liikuntaelinsairauden perusteella määräämäänsä työkyvyttömyyseläkettä, ja vuosia myöhemmin henkilö jää saman psykiatrisen sairauden vuoksi työkyvyttömyyseläkkeelle. Näiden kahden esimerkkihenkilön tapahtumapolku on käytetyssä sekvenssianalyysissä täysin samanlainen. Tämänkaltaisiin esiin nousseihin ilmiöihin ja syy-seuraussuhteisiin on syytä paneutua jatkotutkimuksessa.

Tämän maisterintutkielman tuloksia voidaan kuitenkin pitää aineiston kattavuuden vuoksi hyvin yleistettävänä julkisen sektorin työntekijöihin suomessa eri ikäryhmissä niin miesten kuin naistenkin osalta. Tulosten vertailua muihin Pohjoismaihin voi tietysti varauksin pitää sopivana, mutta tulosten yleistämiseen on suhtauduttava varauksella etenkin muuhun ulkomaiseen tutkimukseen verratessa.

## 7 Johtopäätökset

Maisterintutkielmani tavoite oli tarkastella eksploratiivisella sekvenssianalyysillä onko olemassa sellaisia väestöryhmiä, joissa tuki- ja liikuntaelinsairautena alkaneet työkyvyttömyyseläkkeet vaihtuvat ensisijaiselta lääketieteelliseltä diagnoosiltaan psykiatrian sairaudesta johtuvaksi työkyvyttömyyseläkkeeksi työkyvyttömyyden pitkittyessä. Yhteenvedona voidaan todeta, että tällaisia ryhmiä on. Tietyllä osalla julkisen sektorin työntekijöistä ensisijainen lääketieteellinen syy vaihtuu psykiatriseksi. Tätä tapahtuu tyypillisemmin naisilla ja hieman nuoremmilla työkyvyttömyyseläkettä hakevilla, ja vaihtuminen tapahtuu yleensä ensimmäisen kolmen vuoden aikana ensimmäisestä tuki- ja liikuntaelinsairauden perusteella tehdystä työkyvyttömyyseläkeratkaisusta.

Koska maisterintutkielman tarkoitus oli tutkia vähemmän tunnettua ilmiötä avoimesti ja joustavasti, ja ilman tarkasti määriteltyjä hypoteeseja, voidaan sanoa, että tavoitteen mukaisesti uusia aiheita löydettiin. Tässä työssä pyrittiin ymmärtämään ja löytämään uusia näkökulmia, joita ei välttämättä tiedetty ennalta. Jatkotutkimuksen osalta olisi lukuisia esimerkkejä, joita seurata. Tässä luvussa käyn läpi muutamia selkeitä tuloksista nousseita mahdollisuuksia.

Jatkotutkimusta tarvitaan, että tämä löydetty somaattisena sairautena alkanut, mutta psykiatriseksi sairaudeksi vaihtuneiden joukko voidaan paremmin tunnistaa, ja ylimääräistä kärsimystä helpottaa ja kustannuksia pienentää. Aiemman tutkimuksen perusteella tiedetään, että varhainen puuttuminen, jo sairauspoissaolojen osalta, johtaa parempaan lopputulokseen työkyvyttömyyden kannalta. Kun tunnistetaan paremmin näitä erityisryhmiä, niin voidaan myös paremmin kohdentaa toimenpiteitä niitä tarvitseville. Olisi esimerkiksi mielekästä lisätä muuttujiksi ammattiluokat, sosioekonominen tausta ja koulutustausta, sekä tarkastella missä määrin nämä ovat yhteydessä tautivaihtuvuuteen. Myös erityisesti nuorempia ikäryhmiä voisi tarkastella erikseen sekvenssianalyysin avulla ja tutkia missä määrin vastaavia hakemuspolkuja esiintyy.

Menetelmien osalta jatkotutkimuksen kannalta olisi perusteltua kokeilla erilaisia visualisointimenetelmiä, jotka voisivat tarjota lisäymmärrystä ryhmien sisäisistä rakenteista. Voisi olla hyvä kokeilla erilaisia etäisyyslaskentamenetelmiä (esimerkiksi LCS tai Hamming) ja verrata tuloksia OM-menetelmällä saatuihin tuloksiin. Myös

kustannusparametrejä voisi tarkastella eri substituutiomatriisien (esimerkiksi "CONSTANT") kautta ja arvioida, parantavatko ne analyysin laatua. Eräs konkreettinen vaihtoehto jo tällä samalla aineistolla olisi poistaa ensimmäinen ratkaisu, ja seurata jatkotutkimuksen näkökulmasta muuttuvatko tulokset, kun kaikkein suurin ryhmä (ryhmä 1, yhden ratkaisun jälkeen päättyneet) poistetaan. Lisäksi on huomioitava, että sekvenssianalyysin puuttuvat arvot ovat tunnistettu haaste, ja jatkossa tähän on tehtävä tarkentavaa tutkimusta, vaikkei se tämän aineiston kanssa ollutkaan oleellinen ongelma.

Myös aineiston käsittely voi aineistojen koon kasvaessa, ja erityisesti niin sanotun "big datan" tullessa käyttöön, osoittautua mielenkiintoiseksi haasteeksi. Datan kerääminen edellyttää oman osaamisensa, samoin sen jalostaminen, esimerkiksi koneoppimisen ja syväoppimisen keinoin. On hyvin mahdollista, että datajalostamot nähdään tulevaisuudessa yhtenä kansallisen vaurauden lähteenä. Sekvenssianalyysi voisi hyötyä tällaisista massiivista data-aineistoista, ja voisi olla mielenkiintoista tutkia mitä näitä elementtejä yhdistämällä olisi mahdollista löytää.

Tässäkin työssä käytettyä kluterianalyysia on hyödynnetty esimerkiksi ruotsalaisessa työkyvyttömyyseläketutkimuksessa, jossa havaittiin, että miehet ja naiset eroavat jonkin verran sosioekonomisen aseman suhteen, mutta myös alueellisesti. Tämä viittaa siihen, että Suomessakin olisi perusteltua tutkia alueellisia eroja työkyvyttömyyseläkkeen saajien keskuudessa liittämällä aineistoon koti- tai työpaikkakuntia.

Lisäksi on paljon ammatteja, jotka ovat henkisesti kuormittavia, niissä on matalampi palkkaus ja tehokkuusvaatimukset ovat suuret ja mahdollisuudet vaikuttaa omaan työhönsä ovat pienet. On paljon ihmissuhdeosaamiseen perustuvia naisvaltaisia ammatteja, joissa esiintyy mielenterveysongelmia. Kaikenlaisessa asiantuntijatyössä on nykyään koventuneet paineet, ja eletään työympäristössä, joka on muutoksiltaan nopea, epävarma ja monimutkainen.

Vaikka tämän maisterintutkielman tulokset osoittavat, että tautivaihtuvuutta on hieman enemmän nuoremmissa ikäryhmissä ja naisilla, niin on tärkeä korostaa, että mielenterveyshaasteiden sukupuolittunut rooli voi liittyä iän lisäksi myös työelämään ja laajemminkin siihen minkälaisia toimijoita olemme sekä minkälaisiin identiteetteihin sitoudumme. Sukupuolittuneet olemisen tavat ja normit määrittävät

paljon sitä minkälaisiin ammatteihin ohjaututaan ja toisaalta sitä, miten terveysongelmiin reagoidaan. Kyse ei siis yksioikoisesti ole siitä, että naiset tai nuoremmat ikäryhmät olisivat psyykkisesti haavoittuvampia tai heikommassa asemassa. Tämän maisterintutkielman, ja työkyvyttömyyseläketutkimuksen laajemminkin, lopullinen tarkoitus on parempi ymmärrys työkyvyttömyyseläkkeistä, hyvinvoinnin lisääminen ja työkyvyttömyyden ehkäisy.

## Lähteet

Abbott, A. & Tsay, A. (2000). Sequence analysis and optimal matching methods in sociology. *Sociological Methods & Research*, 29(1), 3-33.

Aisenbrey, S. & Fasang, A. (2010). New life for old ideas: The “second wave” of sequence analysis bringing the “course” back into the life course. *Sociological Methods & Research*, 38(3), 420-462.

Angrist, J. D., & Pischke, J.-S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives*, 24 (2), 3–30.

<https://pubs.aeaweb.org/doi/pdf/10.1257/jep.24.2.3>

Bozio, A., Garrouste, C., & Perdrix, E. (2021). Impact of later retirement on mortality: Evidence from France. *Health Economics*, 30(5), 1178–119.

Brzinsky-Fay, C., U. Kohler, and M. Luniak (2006). Sequence analysis with Stata. *The Stata Journal* 6 (4), 435–460.

Brzinsky-Fay, C. (2007). Lost in transition? Labour market entry sequences of school leavers in Europe. *European Sociological Review*, 23(4), 409-422.

Brzinsky-Fay, C. (2014). Graphical representation of transitions and sequences. *Advances in Sequence Analysis: Theory, Method, Applications*, Cham/Heidelberg/New York/Dordrecht/London: *Springer*. 265-284

Carone, G., Eckefeldt, P., Giamboni, L., Laine, V., & Pamies Sumner, S. (2016). Pension reforms in the EU since the early 2000’s: Achievements and challenges ahead. *DP European Economy*, (042), 1–64.

Compton, M. T., & Shim, R. S. (2015). The social determinants of mental health. *Focus*, 13(4), 419–425.

Elzinga, C. and A. Liefbroer (2007). De-standardization of family-life trajectories of young adults: A cross-national comparison using sequence analysis. *European Journal of Population/Revue européenne de Démographie* 23 (3), 225–250.

Elzinga, C. H. (2008). Sequence analysis: Metric representations of categorical time series. *Sociological Methods and Research*. Forthcoming

ETK Työkyvyttömyyseläkkeelle. Eläketurvakeskus (2024). Haettu 20.10.2024 osoitteesta: <https://www.etk.fi/ajankohtaista/lahes-18-000-jai-tyokyvyttomyyselakkeelle/>

ETK Osittainen vanhuuseläke. Eläketurvakeskus (2024). Haettu 28.10.2024 osoitteesta: <https://www.etk.fi/suomen-elakejarjestelma/elaketurva/tyoelake-etuudet/osittainen-vanhuuselake/>

ETK 2022, Eläketurvakeskus (2022). Haettu 10.7.2024 osoitteesta: <https://www.etk.fi/ajankohtaista/tyokyvyttomyyselakkeelle-siirtyi-ennatyksellisen-vahan-ihmisia/>

ETK Eläkejärjestelmä. Eläketurvakeskus (2024). Haettu 10.2.2024 osoitteesta: <https://www.etk.fi/suomen-elakejarjestelma/elakeuudistukset/aiemmat-uudistukset/>

ETK Työkyvyttömyyseläke. Eläketurvakeskus (2024). Haettu 27.2.2024 osoitteesta <https://www.etk.fi/suomen-elakejarjestelma/elaketurva/tyoelake-etuudet/tyokyvyttomyyselake/>

Ervasti J, Pietiläinen O, Rahkonen O, Lahelma E, Kouvonen A, Lallukka T, Mänty M. (2018). Long term exposure to heavy physical work, disability pension due to musculoskeletal disorders and all-cause mortality—20-year follow-up: introducing Helsinki Health Study job exposure matrix. *Int Arch Occup Environ Health*.

Ervasti, J., Kouvonen, A., Laaksonen, M., Lahelma, E., Lahti, J., Lallukka, T., ... & Mänty, M. (2019). *Työkuormitus, työ- ja toimintakyky*. Työterveyslaitos: Helsinki

Fasang, A. & Liao, T. (2014). Visualizing sequences in the social sciences. *Sociological Methods & Research*, 43(4), 643-676.

Fleischmann M., Xue B., Head J. (2020). Mental health before and after retirement—assessing the relevance of psychosocial working conditions: the Whitehall II *Prospective Study of British Civil Servants*. 75:403–13.

- Gabadinho, A., G. Ritschard, M. Studer and N. S. Müller, (2009). *Mining sequence data in R with the TraMineR package: A user's guide*, University of Geneva. (<http://traminer.unige.ch>)
- Gabadinho, A., G. Ritschard, N.S. Müller and M. Studer (2011). Analyzing and Visualizing State Sequences in R with TraMineR. *Journal of Statistical Software*, 40(4), 1–37.
- Gabadinho, A., & Ritschard, G. (2013). Searching for typical life trajectories applied to childbirth histories. In Proceedings of the International Conference on Social Science Methodology.
- Galaasen A, Bruusgaard D, Claussen B. (2012). Excluded from social security: Rejections of disability pension applications in Norway 1998–2004. *Scand J Public Health* 2012; 40: 142–149.
- Garrouste, C., & Perdrix, E. (2022). Is there a consensus on the health consequences of retirement? A literature review. *Journal of Economic Surveys*, 36: 841–879).
- Gold S., Köhler-Forsberg O., Moss-Morris R., Mehnert, A., Miranda, J., Bullinger, M., Steptoe, A., Whooley, M. & Otte, C. (2020). Comorbid depression in medical diseases. *Nat Rev Dis Primers* 2020; 6:69.
- Goodell, S., Druss, B. G. & Walker, E. R. (2011). Mental disorders and medical comorbidity. *Robert Wood Johnson Foundation*. The synthesis project, No.21, 2011.
- Halpin, B. (2016a). Missingness and truncation in sequence data: A non-self-identical missing state. *Proceedings of the international Conference on Sequence Analysis and Related Methods*. (s.443-444) Lausanne.
- Halpin, B (2016b). Multiple imputation for categorical time series. *Stata Journal*, 17(3), s.546–612.
- Heikkala E., Oura P., Ruokolainen O., Ala-Mursula L., Linton, S. & Karppinen, J. (2023). The Örebro Musculoskeletal Pain Screening Questionnaire-Short Form and 2-year follow-up of registered work disability. *Eur J Public Health* 2023; 33: 442–7.
- Heinonen, N., Lallukka, T., Lahti, J., Pietiläinen, O., Nordquist, H., Mänty, M., Katainen, A., Kouvonen, A. (2022). Working Conditions and Long-Term Sickness

Absence Due to Mental Disorders: A Prospective Record Linkage Cohort Study Among 19- to 39-Year-Old Female Municipal Employees. NIH, *National Library of Medicine*, 1;64(2):105–114.

Helgesson, M., Rahman, S., Saboonchi, F., Mittendorfer Rutz, E. (2021). Disability pension and mortality in individuals with specific somatic and mental disorders: examining differences between refugees and Swedish-born individuals. *J Epidemiol Community Health*. 2021 Aug;75(8):721-728. doi: 10.1136/jech-2019-213436. Epub 2021 Jan 20.

Henkilötietolaki 22.4.1999/523.

Kannisto, J. (2022) Muuttuva vanhuuseläkeikä: ikärajamuutosten seurantatilasto 2021. *Eläketurvakeskuksen tilastoja 12/2022*: Helsinki.

KELA Työkyvyttömyyseläke. Kansaneläkelaitos (2024). Työkyvyttömyyseläkettä saaneiden nuorten määrän väheneminen jatkui vuonna 2022. Päivitetty 13.9.2024.

Haettu 20.10.2024 osoitteesta:

<https://tietotarjotin.fi/uutinen/779315/tyokyvyttomyyselaketta-saaneiden-nuorten-maaran-vaheneminen-jatkui-vuonna-2022>

KELA Kelan historia. Kansaneläkelaitos (2024). Haettu 27.2.2024 osoitteesta:

<https://www.kela.fi/historia#koyhainhoidosta-sosiaaliturvaan-1900-luvulla>

KELA Mielenterveyden häiriöt. Kansaneläkelaitos (2024). Mielenterveyden häiriöistä johtuvien sairauspoissaolojen kasvu jatkuu – kehityksen taustalla useita yhtä aikaa vaikuttavia tekijöitä. Haettu 18.7.2024 osoitteesta:

<https://tietotarjotin.fi/tutkimusblogi/1014063/mielenterveyden-hairioista-johtuvien-sairauspoissaolojen-kasvu-jatkuu-kehityksen-taustalla-useita-yhta-aikaa-vaikuttavia-tekijoita>

Kouvonen A, Mäntä M, Lallukka T, Lahelma E, Rahkonen O. (2016). Changes in psychosocial and physical working conditions and common mental disorders. *Eur J Public Health* ;26(3):458–463.

Kouvonen A, Mäntä M, Lallukka T, Pietiläinen O, Lahelma E, Rahkonen O. (2017). Changes in psychosocial and physical working conditions and psychotropic

medication in ageing public sector employees: a record-linkage follow-up study. *BMJ Open*.

Laaksonen, M. (2016) Koetun työkyvyn ja pitkäaikaissairastavuuden muutokset sosioekonomisten tekijöiden ja työolojen mukaan. Teoksessa Noora Järnefelt (toim.) *Työolot ja työurat – tutkimuksia työurien vakaudesta ja eläkkeelle siirtymisestä*. Helsinki: Eläketurvakeskus, 107–124.

Laaksonen, M. (2020) Työkyvyttömyyseläkkeelle siirtyminen: trendit, tilannekuva, tulevaisuus. *Eläketurvakeskuksen raportteja 07/2020*. Eläketurvakeskus: Helsinki.

Laaksonen, M. & Elovainio, M. & Kainulainen, S. & Leinonen, T. & Jääskeläinen, T. & Rissanen, H. & Koskinen, S. (2022) Changes in healthy and unhealthy working life expectancies among older working-age people in Finland, 2000–2017. *European Journal of Public Health*, 32(5): 729–734.

Laki Kevasta 66/2016. Haettu 10.7.2024 osoitteesta:

<https://www.finlex.fi/fi/laki/ajantasa/2016/20160066>

Lappo, S. ja Risku, I. (2023). *Kansantaloudellinen aikakauskirja 2/2023*, ”Työkyvyttömyyseläkkeiden kustannus”. Haettu 14.10.2023 osoitteesta

<https://urn.fi/URN:NBN:fi-fe20230921134883>

Lesnard, Laurent (2009). “Cost Setting in Optimal Matching to Uncover Contemporaneous Socio-Temporal Patterns”, *Notes & Documents*, 2009-03, Paris, OSC, Sciences Po/CNRS.

Leppävuori A. (2023) Depressio ja somaattinen sairaus. *Duodecim* 2023; 139:1367–76.

Liao, T., Bolano, D., Brzinsky-Fay, C., Cornwell, B., Fasang, A., Helske, S., Piccarreta, R., Raab, M., Ritschard, G., Struffolino, E., & Studer, M. (2022). Sequence analysis: Its past, present, and future. *Social Science Research*, Volume 107,

<https://doi.org/10.1016/j.ssresearch.2022.102772>.

Mathisen, J., Koitzsch Jensen, N., Bjorner, J., Brønnum-Hansen, H., Christensen, U., Thielen, K., Gadeberg, A., Andersen, I. (2021). Disability pension among persons with chronic disease: Differential impact of a Danish policy reform, *European Journal of*

*Public Health*, Volume 31, Issue 1, February 2021, Pages 186–192,

<https://doi.org/10.1093/eurpub/ckaa228>

Mänty M, Kouvonen A, Lallukka T, Lahti J, Lahelma E, Rahkonen O. (2015). Changes in working conditions and physical health functioning among midlife and ageing employees. *Scand J Work Environ Health*; 41(6):511–518.

Mänty M, Kouvonen A, Lahti J, Lahelma E, Rahkonen O. (2018). Changes in physical and mental health functioning during retirement transition: a register-linkage follow-up study. *Eur J Public Health*; 28(5):805–809.

Mänty M, Kouvonen A, Lallukka T, Lahti J, Lahelma E, Rahkonen O. (2016). Pre-retirement physical working conditions and changes in physical health functioning during retirement transition process. *Scand J Work Environ Health*; 42(5):405–412.

Mänty M, Lallukka T, Lahti J, Pietiläinen O, Laaksonen M, Lahelma E, Rahkonen O. (2017). Physical and mental health functioning after all-cause and diagnosis specific sickness absence: a register-linkage follow-up study among ageing employees. *BMC Public Health*; 17:114–122.

Mänty, M., Kouvonen, A., Nordquist, H., Harkko, J., Pietiläinen, O., Halonen, J. I., Rahkonen, O., ja Lallukka, T. (2022). Physical working conditions and subsequent sickness absence: a record linkage follow-up study among 19–39-year-old municipal employees. *Int Arch Occup Environ Health*. 2022; 95(2): sivut 489–497.

Nishimura, Y., Oikawa, M., & Motegi, H. (2018). What explains the difference in the effect of retirement on health? Evidence from global aging data. *Journal of Economic Surveys*, 32(3), 792–847.

Nivalainen, S., Ilmakunnas, I., Laaksonen, M. (2023). Tutkimuksia vanhuus- ja työkyvyttömyyseläkkeelle siirtymisen muutoksista vuoden 2017 eläkeuudistuksen jälkeen. *Eläketurvakeskuksen tutkimuksia* 1/2023. Haettu 13.11.2023 osoitteesta: <https://www.julkari.fi/bitstream/handle/10024/145972/tutkimuksia-vanhuus-ja-tyokyvyttomyyselakkeelle-siirtymisen-muutoksista-vuoden-2017-elakeuudistuksen-jalkeen.pdf?sequence=12&isAllowed=y>

- Perhoniemi, R., Blomgren, J., Laaksonen, M. (2020). Determinants of disability pension applications and awarded disability pensions in Finland, 2009 and 2014. *Scandinavian Journal of Public Health*. 2020;48(2):172–180.
- Raab, M. & Struffolino, E. (2023). *Sequence Analysis*. Sage: California.
- Rissanen, K. ja Kaseva E. (2014). ”Menetetyn työpanoksen kustannus”. *Sosiaali- ja terveystieteiden tutkimuskeskus*. Haettu 14.10.2023 osoitteesta: [https://stm.fi/documents/1271139/1332445/Menetetyn+ty%C3%B6panoksen+kustannus+2+\(2\)+\(2\).pdf](https://stm.fi/documents/1271139/1332445/Menetetyn+ty%C3%B6panoksen+kustannus+2+(2)+(2).pdf)
- Rohwer, G. and U. Pötter (2002). *TDA user's manual*. Software, Ruhr-Universität Bochum, Fakultät für Sozialwissenschaften, Bochum.
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53-65.
- Ruponen, P. (2023). Predicting diagnosis classes from medical text using deep transformer-based models. Aalto University, master's thesis.
- Ryynänen, K. & Metsä, T. (2024). Lisääkö tuki- ja liikuntaelinsairaus mielenterveysoireita? *Lääkärilehti* 79(17–18), 700–703.
- STM, Sosiaali- ja terveystieteiden tutkimuskeskus (2017). Työryhmäraportti – Kannustinloukut. Työkyvyttömyyseläkkeen ja ansiotulojen yhteensovittaminen sekä lineaarinen malli. *Osatyökykyisille tie työelämään –kärkihanke* (OTE). Haettu 11.12.2023 osoitteesta: [https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/80691/RAP2017\\_32.pdf?sequence=1&isAllowed=y](https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/80691/RAP2017_32.pdf?sequence=1&isAllowed=y)
- Studer, M. ja Ritschard, G. (2016). What Matters in Differences Between Life Trajectories: A Comparative Review of Sequence Dissimilarity Measures. *Journal of the Royal Statistical Society Series A (Statistics in Society)*. 179. 481–511. 10.1111/rssa.12125.
- Tutkimuseettinen neuvottelukunta (TENK). (2023). *Hyvä tieteellinen käytäntö ja sen loukkausepäilyjen käsittelyminen Suomessa*. Tutkimuseettisen neuvottelukunnan HTK-ohje. Tutkimuseettisen neuvottelukunnan julkaisuja 2/2023. Helsinki.

THL ICD-10 (2011). Julkari. Haettu 11.12.2023 osoitteesta:

<https://urn.fi/URN:NBN:fi-fe201205085423>

Tekemättömän työn vuosikatsaus 2016, Yhteenveto yritysaineiston tuloksista (2016).

Haettu 14.10.2023 osoitteesta:

[https://www.keva.fi/globalassets/2-tiedostot/tama-on-keva--tiedostot/tutkimukset/tutkimus\\_tekemattoman-tyon-vuosikatsaus-2016-tulosraportti.pdf](https://www.keva.fi/globalassets/2-tiedostot/tama-on-keva--tiedostot/tutkimukset/tutkimus_tekemattoman-tyon-vuosikatsaus-2016-tulosraportti.pdf)

Telma (2023). ”Kuinka kalliiksi työkyvyttömyys tulee?”. Haettu 14.10.2023

osoitteesta <https://telma-lehti.fi/kuinka-kalliiksi-tyokyvyttomyys-tulee>

Valtioneuvosto (2023). ”Vuosien 2024–2027 tekninen julkisen talouden suunnitelma

julkaistu”. Haettu 14.10.2023 osoitteesta <https://valtioneuvosto.fi/-/10616/vuosien-2024-2027-tekninen-julkisen-talouden-suunnitelma-julkaistu>

Virtanen, M., Heikkilä, K., Vahtera, J., Kivimäki, M., Halonen, JI., Alexanderson, K.,

Rautiainen, S., Lallukka, T., Mittendorfer-Rutz, E. (2022). Clustering of disability

pension and socioeconomic disadvantage in Sweden: a geospatial analysis. *Eur J*

*Public Health*. 2022 Oct 3;32(5):703-708.

## Liitteet

### Liite 1. Tutkimuksessa ja tutkielmassa käytetyt ohjelmat/ohjelmistot

OpenAI. (2024). ChatGPT (GPT-4-turbo versio) [Kielimalli]. Käytetty R-koodin oikolukemiseen ja tarkistamiseen 1.6.-31.10.2024 välillä osoitteessa:

<https://chat.openai.com/chat>

Käyttö Helsingin yliopistossa: <https://studies.helsinki.fi/ohjeet/artikkeli/tekoalyn-kayttaminen-oppimisen-tukena>

R-ohjelmisto (2024). R: A language and environment for statistical computing (version 4.3.3). <https://www.r-project.org/>