



Noora Siistonen

Tekoälyn hyödyntäminen diabeettisen retinopatian seulontakuvauksessa

Scoping -kirjallisuuskatsaus

Metropolia Ammattikorkeakoulu

Optometristi YAMK

Kliininen asiantuntijuus, digitaalisten palveluiden asiantuntija

Opinnäytetyö 18.12.2024

Tiivistelmä

Tekijä:	Noora Siistonen
Otsikko:	Tekoälyn hyödyntäminen diabeettisen retinopatian seulontakuvauksessa, Scoping -kirjallisuuskatsaus
Sivumäärä:	53 sivua
Aika:	18.12.2024
Tutkinto:	Optometrismi (ylempi AMK)
Tutkinto-ohjelma:	Kliininen asiantuntijuus, digitaalisten palveluiden asiantuntija
Ohjaajat:	Yliopettaja Anu Valtonen Lehtori Saila Pakarinen

Diabetesdiagnoosien kasvu lisää painetta tuottaa riittävästi silmänpohjan kuvaseulontaa ja lausuntoja. Väestörakenteen muutos ja huoltosuhteen heikentyminen tulevat lisäksi vaikuttamaan tulevaisuudessa sote-ammattilaisten rekryointivaikeuksiin. Yhtenä ratkaisuna voi olla palveluiden tehostaminen tekoälyn avulla. Tämän opinnäytetyön tarkoituksena oli selvittää, miten tekoälyä voidaan hyödyntää diabeettisen retinopatian vaikeusasteen luokittelussa. Tavoitteena oli tuottaa tietoa tekoälyn luotettavuudesta ja tarkkuudesta luokittelussa ja sen käytön hyödyistä ja haitoista. Menetelmänä käytettiin kartoitettavaa scoping -kirjallisuuskatsausta. Tietokantahaku tehtiin Cinahl, Taylor & Francis Online, Sage Journals sekä Medic tietokannoissa. Etukäteen määriteltyjen sisäaiontokeriteerien mukaisia tutkimusartikkeleita löytyi 18, joille tehtiin laadunarviointi ja aineistolähtöinen sisällönanalyysi.

Tekoälyalgoritmin käytöstä voi olla useita hyötyjä: Sillä saavutetaan kustannustehokkuutta ja kustannussäästöjä, kun herkkyys, tarkkuus ja luokittelun kynnyksarvot on viilattu sopimaan mm. maan tautitilanteeseen. Tekoälyalgoritmi pystyy parhaimmillaan lausumaan automaattisesti, luotettavasti ja nopeasti jokaisen sille syötetyn kuvan. Se voi vaikuttaa tuottavuuteen säästämällä silmälääkäreiden aikaa ja resursseja vaativimpiin tapauksiin. Lisäksi algoritmi voi auttaa kokemattomia kuvaajia varoittamalla riittämättömästä kuvalaadusta tai lisäkuvien tarpeesta. Haittapuolina algoritmien suorituskyky voi vaihdella ja heikoimpien algoritmien käyttö ei olisi kustannustehokasta tai turvallista potilaan kannalta ilman laadunvalvontaa. Joiltain algoritmeilta jäi myös suuri osa kuvista lausumatta, mikä lisäisi silmälääkäreiden työmäärää. Lisäksi saman algoritmin tutkimustuloksissa oli hajontaa.

Yhteenvedona voidaan todeta, että parhaimmillaan tekoälyalgoritmi voi olla tarkka ja luotettava, jolloin sen käytöstä seuraa hyötyjä. Tekoäly voi tehostaa diabeettisen retinopatian seulontaa ja vastata näin kasvavaan kuvantamistarpeeseen. Algoritmin valinnassa ja hankinnassa on kuitenkin oltava tarkkana, jotta hyödyt maksimoidaan ja haitat minimoidaan.

Avainsanat: diabeettinen retinopatia, tekoäly

Tämän opinnäytetyön alkuperä on tarkastettu Turnitin Originality Check -ohjelmalla

Abstract

Author: Noora Siistonen
Title: The Utilization of Artificial Intelligence in Diabetic Retinopathy Screening, a Scoping Literature Review
Number of Pages: 53 pages
Date: 18 December 2024

Degree: Master of Health Care (Optometry)
Degree Programme: Master's Degree Programme in Clinical Expertise in Digital Health Care and Social Services
Instructors: Anu Valtonen, Principal Lecturer
Saila Pakarinen, Senior Lecturer

The increase in diabetes diagnoses adds pressure to produce sufficient retinal imaging screenings and statements. Changes in population structure and a weakening dependency ratio will also affect the recruitment difficulties of healthcare professionals in the future. Enhancing services with artificial intelligence could be a solution. The purpose of this thesis was to investigate how artificial intelligence can be utilized in classifying the severity of diabetic retinopathy. The aim was to provide information on whether artificial intelligence is reliable and accurate enough in classification and what benefits or drawbacks have been observed from its use. The method used was a scoping literature review. Database searches were conducted in Cinahl, Taylor & Francis Online, Sage Journals and Medic databases. 18 research articles meeting the predefined inclusion criteria were selected to quality assessment and data-driven content analysis.

There can be several benefits to using artificial intelligence: cost-effectiveness and cost savings when their sensitivity, accuracy and classification thresholds are fine-tuned to suit, among other factors, the disease prevalence in the country. At their best, artificial intelligence algorithms can automatically, reliably, and quickly provide classification for every image. This can enhance productivity by saving ophthalmologists' time and resources for more demanding cases. Additionally, artificial intelligence algorithms can assist less experienced image capturers by alerting them to insufficient image quality or the need for additional images. However, the performance of algorithms can vary, and using the least effective ones would not be cost-effective or safe for the patient without quality control. Some algorithms also failed to classify certain images, which would increase the workload for ophthalmologists. Furthermore, there was variation in the results of studies evaluating the same algorithm.

In summary, it can be stated that at best, an artificial intelligence algorithm can be accurate and reliable resulting in benefits from its use. Artificial intelligence can help improve the efficiency of diabetic retinopathy screening, thereby meeting the growing demand for imaging. However, careful consideration is needed in the selection and acquisition of the algorithm to maximize benefits and minimize drawbacks.

Keywords: diabetic retinopathy, artificial intelligence, AI

The originality of this thesis has been checked using Turnitin Originality Check service.

Sisällys

1	Johdanto	1
2	Teoreettinen viitekehys	2
2.1	Diabetes	2
2.2	Diabeettinen retinopatia	3
2.2.1	Diabeettisen retinopatian luokittelu	4
2.2.2	Seulontakuvaus silmnpohjakuvilla	6
2.2.3	Seulontakuvien lausuminen	7
2.2.4	Diagnostisen testin herkkyys ja tarkkuus	8
2.3	Tekoäly	9
2.3.1	Tekoäly terveydenhuollossa	9
2.3.2	Tekoäly diabeettisen retinopatian seulontakuvauksessa	10
3	Kirjallisuuskatsauksen tarkoitus ja tavoitteet	11
4	Kirjallisuuskatsauksen toteutus	12
4.1	Tiedonhaun kuvaus	13
4.1.1	Tietokannat	13
4.1.2	Hakusanat	13
4.1.3	Sisäänotto- ja poissulkukriteerit	15
4.1.4	PRISMA –kaavio	17
4.2	Aineiston laadun arviointi	18
4.3	Aineiston analysointi	19
4.4	Tutkimusten yhteenveto	26
5	Tulokset	28
5.1	Tekoälyn käytön hyötyjä diabeettisessa retinopatiassa	28
5.1.1	Tekoälyn suoriutuminen	29
5.1.2	Tekoälyn hyviä ominaisuuksia	30
5.1.3	Tekoälyn käytön seurauksia	31
5.2	Tekoälyn käytön haittoja diabeettisessa retinopatiassa	32
5.2.1	Tekoälyn suoriutuminen	33

5.2.2	Tekoälyn ominaisuudet	37
6	Pohdinta	37
6.1	Tulosten yhteenveto	37
6.2	Päätulosten pohdinta	38
6.3	Hyödynnettävyys ja jatkotutkimusaiheet	43
6.4	Eettisyys ja luotettavuus	44
	Lähteet	48

1 Johdanto

Diabetesta sairastavia on Suomessa arviolta 500 000 (Diabetesliitto 2021). Kattavaa tilastointia diagnoosimääristä ei ole, mutta vuosien saatossa diabetesta sairastavien määrä on kasvanut tasaisesti n. 20 000 vuodessa (Ilanne-Parikka 2021). Diabeettinen retinopatia kuuluu diabeteksen lisäsairauksiin (THL 2023), ja sen seulonta kuuluu diabetesta sairastavan hoitoon (Seppänen 2021). Diabeteksen hoito maksaa noin 1300 euroa vuodessa potilasta kohden, mutta eri lisäsairauksien kanssa summa on noin 5700 euroa vuodessa (THL 2023). Varhainen puuttuminen lisäsairauksiin vähentää taudin haittoja potilaalle. Kuten diabeettisessa retinopatiassa on uhka merkittäväälle näön heikentymiselle. Oikein ajoitettuna hoito auttaa säilyttämään näkökykyä. (Seppänen 2021.) Varhainen puuttuminen myös vähentää kustannuksia yhteiskunnalle (Von Wendt & Komulainen 2014). Pelkästään diagnoosimäärien kasvu lisää kuormaa riittäväälle seulontakuvauksen määrälle ja edelleen taakkaa silmätautien erikoislääkäreille kasvavina lausuntomäärinä (Sahlsten & Jaskari & Kaski & Hietala 2020).

Väestörakenne Suomessa on muuttumassa: syntyvyys laskee, työkäisten määrä vähenee ja väestö ikääntyy ja elää pidempään (Jalovaara ym. 2023: 2, 4). Huoltosuhteen heikkeneminen aiheuttaa sen, että tulevia terveydenhuollon ammattilaisia ei tule valmistumaan riittävää määrää terveydenhuollon tarpeisiin nähden, sillä nuoria tulee kouluttaa myös muille aloille. Rekrytointivaikeudet tulevat kasvamaan kaikkien sote-ammattiryhmien osalta. Henkilökunnan tarpeen ja tarjonnan välinen epäsuhta on kasvamassa sekä yksityisellä, että julkisella puolella. (Lehtonen & Saari & Kinnunen & Kinnunen 2023: 132–134.) Väestön vanheneminen tulee lisäämään kysyntää myös silmä lääketieteen palveluille, sillä moni silmätauti yleistyy (Hautala 2020).

Helsingin ja Uudenmaan sairaanhoitopiirissä marraskuun 2024 lopussa kaikista 16 erikoisalasta silmätaudeilla oli kolmanneksi pisin hoitopääsyt mediaaniaika, 67 päivää (HUS 2024). Väestöllisen ja taloudellisen huoltosuhteen heikkeneminen tulee lisäämään painetta sosiaali- ja terveyspalveluiden tuottamiseen hyvinvointivaltiossa. Yhtenä ratkaisuna on palveluiden tehostaminen. (Schleutker 2013.) Palveluiden tehostamisen yhtenä vaihtoehtona on pidetty teknologisia ratkaisuja ja tekoälyyn pohjautuvien sovelluksien hyödyntämistä sote-työssä (Kirkonpelto & Mäntyranta 2023). Esimerkiksi radiologiassa tekoälyä käytetään jo tulkitsemaan mammografian seulontakuvia kasvainten havaitsemisessa, jolloin kahden radiologin sijaan lausunnon tuottaa yksi radiologi tekoälyn avustamana (Huhtanen & Nyman & Karlsson & Hirvonen 2020). Tekoälyn tekemät

diabeettisen retinopatian luokittelut on Suomessa tehdyn tutkimuksen mukaan todettu olevan riittävän herkkiä ja tarkkoja, jotta seulonnan laaturaja täyttyy (Sahlsten ym. 2020). Tekoäly on parhaimmillaan rajatussa tehtäväasetelmassa ja toistuvassa päätöksenteossa eli esimerkiksi kuvantamisessa (Kolari & Kallio 2023: 16–17).

Tämän opinnäytetyön tarkoituksena on selvittää scoping -kirjallisuuskatsauksen keinoin, miten tekoälyä pystytään hyödyntämään diabeettisen retinopatian vaikeusasteen luokittelussa. Opinnäytetyön tavoitteena on tuottaa tietoa, voiko tekoälyä käyttää kuvien luokittelussa korvaamaan osittain silmälääkärin, tai muun lausujan tekemät lausunnot. Tavoitteena on myös tuottaa tietoa, onko tekoäly tarpeeksi luotettava ja tarkka analyyseissaan, sekä mitä hyötyä tai haittaa tekoälyn käytöstä on ilmennyt.

2 Teoreettinen viitekehys

Tiedonhakuja teoreettisen viitekehysten ja johdannon kokoamiseksi suoritettiin manuaalisesti luotettaviksi tiedetyistä internetlähteistä. Lisäksi käytettiin hyödyksi alan julkaisuja, aikaisempia tutkimuksia ja käytiin läpi niiden lähdeviitteitä. Myös kirjastosta lainattiin sopivia aiheisiin liittyviä kirjoja nimen, takakannen ja sisällysluettelon perusteella. Lisäksi suoritettiin tietokantahakuja asiasanoilla muutamista tietokannoista. Metropolia Ammattikorkeakoulun MetCat FINNASTA hakua tehtiin muun muassa termeillä “artificial intelligence” AND (medicine OR healthcare) sekä “artificial intelligence” AND imaging. Suomalaisesta terveysalan artikkeleita sisältävästä Medic tietokannasta etsittiin tietoa hakusanalla tekoäly.

2.1 Diabetes

Diabetes on yhteisnimitys taustaltaan erilaisille sairauksille, joiden yhdistävä tekijä on koholla oleva verensokeri. Verensokerin kohonnut pitoisuus voi johtua insuliinihormonin puuttumisesta tai ettei se vaikuta elimistössä oikein tai näiden yhteisvaikutuksesta. Diabetestyyppinä on luokiteltu eri luokkiin, joista yleisimmät ovat tyypin 1 (10–15 % osuus diagnooseista) ja tyypin 2 (75–80 % osuus diagnooseista) diabetekset. Muita diabetestyyppinä on muun muassa raskausdiabetes, LADA ja MODY. (THL 2023.) Diabetes voi myös muuttua ajan saatossa, joten tärkeintä on hoitaa sairautta vaikeusasteen mukaisesti ja näin ehkäistä eri komplikaatioita, joita diabetes aiheuttaa (Diabetes. Käypä hoito –suositus 2018).

Diabetes on suomalaisten kansantauti, sairastavien määrä on yli 500 000 kasvaen koko ajan (Diabetes. Käypä hoito –suositus 2018; Ilanne-Parikka 2021). Vuosittain uusia diagnooseja Suomessa todetaan noin 20 000 (Ilanne-Parikka 2021). Diabetesta sairastaa globaalisti arviolta 422 miljoonaa ihmistä ja on huomattu, että edeltävien vuosikymmenten aikana diagnoosimäärät ovat kasvaneet tasaisesti. Varsinkin 2 tyypin diabetes painottuu matalan- ja keskitulon maihin. (WHO 2024.) Kansainvälinen diabetesliitto ennustaa diabetesta sairastavien määrän nousevan 783 miljoonaan vuoteen 2045 mennessä (IDF 2024).

2.2 Diabeettinen retinopatia

Diabetes aiheuttaa lukuisia komplikaatioita ja lisäsairauksia, varsinkin jos sairautta ei hoideta riittävästi. Eritoten pitkäaikainen altistus suurentuneelle verensokerin pitoisuudelle kasvattaa riskiä lisäsairauksille. Riskiä lisää myös tupakointi, kohonnut kolesteroli ja verenpaine. Yksi merkittävä diabeteksen aiheuttama komplikaatio on diabeettinen retinopatia eli silmän verkkokalvosairaus. (Ilanne-Parikka 2021.) Diabeettista retinopatiaa esiintyy noin 50 %:lla tyypin 1 diabetesta sairastavilla (vähintään lievääasteinen luokittelu), sekä noin 25 %:lla tyypin 2 diabetesta sairastavilla. Taudin kesto lisää todennäköisyyttä diabeettisen retinopatian esiintyvyyteen. Esimerkiksi yli 20 vuoden diagnoosiaika 1 tyypin diabetesta sairastavilla aiheuttaa yli 80 %:lla muutoksia verkkokalvolle. Vakavammat muutokset ovat yleisempiä heillä, jotka ovat saaneet nuorena tyypin 1 diabetesdiagnoosin, jolloin myös tautia sairastetaan pidempään. (Seppänen 2021.)

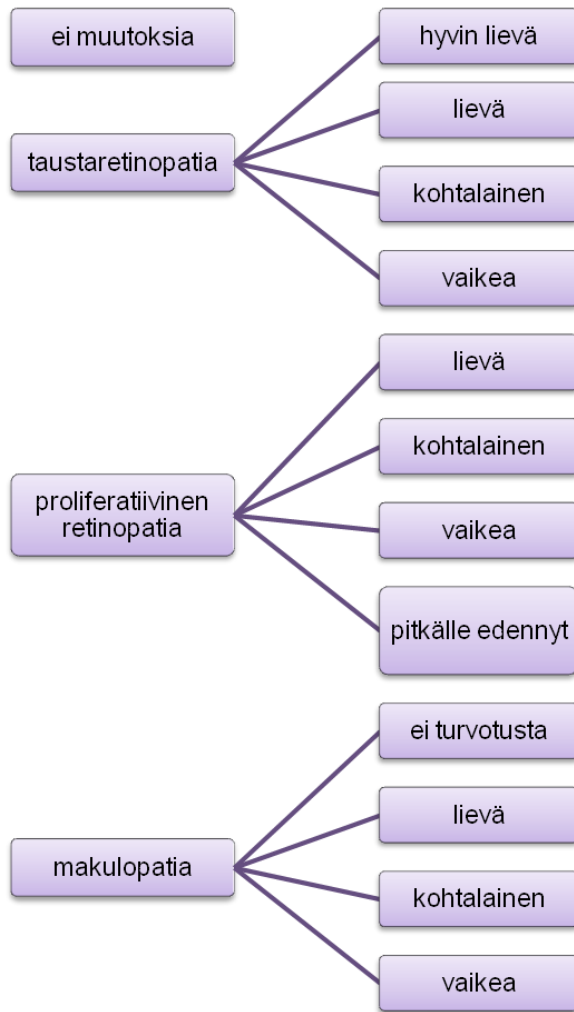
Diabeteksessa veren liian iso sokeripitoisuus vaikuttaa verisuonten sisäpintojen solujen toimintaan, aiheuttaen niissä häiriön. Solujen toimintahäiriö aiheuttaa veren tihkumista verisuonista verkkokalvolle sekä suonet voivat tukkeutua. Tämä puolestaan voi johtaa verkkokalvon hapenpuutteeseen, joka aiheuttaa uudissuonten kasvua. Mikäli hoitotasapaino on pitkään heikko, ja silmänpohjan huonontuneeseen tilanteeseen ei puututa, aiheuttaa diabeettinen retinopatia oireettoman alkuvaiheen jälkeen näön heikkenemistä, näkökenttäpuutoksia ja lopulta tilanne voi edetä sokeuteen. (Seppänen 2021.) Sokeutuminen tapahtuu suurimmalla osalla potilaista 5–10 vuoden aikana, jos proliferatiivista retinopatiaa ei hoida (Diabeettinen retinopatia. Käypä hoito –suositus 2024).

Suomessa diabeteksen aiheuttama sokeutuminen on puolittunut vuodesta 1990 vuoteen 2000. (Diabeettinen retinopatia. Käypä hoito –suositus 2014). Vuoden 2022 loppuun mennessä kaikista rekisteröidyistä näkövammaisista diabetes oli syynä 3,2 %. Tässä kannattaa huomioida, että rekisteri-ilmoitus tehdään sen diagnoosin perusteella, joka on viimeiseksi aiheuttanut paremman silmän näön huonontumista. 3,2 prosentin

osuudella diabeettinen retinopatia on kuudenneksi yleisin syy näkövamman aiheuttajana. Vuoden aikana tapahtuneet uudet näkövammaiset tilastoidaan erikseen. Viimeisten kahdeksan vuoden aikana taustaretinopatian vuoksi näkövammautuneita on ollut alle yksi prosentti vuodessa. Proliferatiivisen retinopatian aiheuttama uusi näkövammaisuus on ollut alle kaksi prosenttia viimeisten viiden vuoden aikana. Huomionarvoista on, että työikäisillä 18–64 –vuotiailla uusilla näkövammaisilla diabeettinen retinopatia on ollut syynä 7,4 prosentilla. (Tolkkinen 2022.)

2.2.1 Diabeettisen retinopatian luokittelu

Diabeettisessa retinopatiassa on käytössä eri asteisia luokitteluja. Luokittelu pyrkii siihen, että sen avulla viestintä on helpompaa lääkäreiden välillä. Luokittelun avulla helpotetaan huomaamaan taudin eteneminen ja oikea-aikaisen päätöksen tekeminen hoidon aloittamiseksi. (Gawęcki 2021:95.) Kuviossa 1 on hahmotelma luokittelusta. Vasemman sarakkeen yläluokittelu tapahtuu neljällä yläotsikolla: ei muutoksia, taustaretinopatia (muutokset eivät uhkaa näkökykyä), proliferatiivinen retinopatia (verkkokalvolla tai näköhermon päässä vähintään yksi uudissuoni) sekä makulopatia (muutokset ovat tarkan näön alueella eli makulassa). Lisäksi nämä kolme yläluokkaa voidaan jakaa edelleen kolmeen tai useampaan alaluokkaan. Esimerkiksi hyvin lievä, lievä, kohtalainen, vaikea, hyvin vaikea ja niin edelleen. (Seppänen 2021.)



Kuvio 1. Diabeettisen retinopatian luokittelu vasemman sarakkeen neljään yläluokkaan ja oikeassa sarakkeessa niiden alaluokkia (mukaillen Diabeettinen retinopatia. Käypä hoito – suositus 2024).

Taustaretinopatissa silmänpohjanmuutokset ovat ensin pieniä pistemäisiä pullistumia hiussuonissa eli mikroaneurysmia, tällöin puhutaan hyvin lievästä diabeettisesta retinopatiasta. Jos mikroaneurysmien lisäksi silmänpohjassa on nähtävillä pieni verenvuotolaikku tai kertynyt rasvaa kiteytyneesti eli muodostunut lipidejä, tai mikroinfarkteja tai venopatiaa puhutaan lievästä diabeettisesta retinopatiasta. Kohtalaiseksi taustaretinopatia muuttuu, jos verenvuotolaikku on iso, tai näkyvissä on IRMAa (intraretinaalinen mikrovaskulaarinen muutos) (Seppänen 2021; Diabeettinen retinopatia. Käypä hoito – suositus 2024.) Proliferatiivisessa retinopatiassa muodostuu uudissuonia verkkokalvon hapenpuutteen seurauksena. Suonet voivat aiheuttaa verenvuotoja verkkokalvolle ja lasiaiseen. Tila voi edetä verkkokalvon irtaumaan, uudissuoniglaukoomaan (vaikeassa hapenpuutteessa) tai suureen lasiaisverenvuotoon. Uudissuoniin voi myös kehittyä fibrovaskulaarista muutosta, eli arpikalvoa, joka saattaa irrottaa verkkokalvoa tai tehdä

reikiä siihen. (Diabeettinen retinopatia. Käypä hoito –suositus 2024.) Makulopatiassa muutokset tapahtuvat tarkan näkemisen alueella, jolloin mahdollinen verkkokalvon turvotus aiheuttaa näön heikentymistä. Makulopatia ei ole riippuvainen diabeettisen retinopatian vaiheesta, vaan sitä voi esiintyä jokaisessa vaiheessa. Makulan turvotusta tulee n. 20 %:lle 2 tyypin diabetesta sairastaville, kun sairastettuja vuosia on takana kymmenen. (Tarnanen & Summanen & Komulainen 2024.) Makulopatiassa luokittelu eri vaikeusasteisiin riippuu siitä, kuinka kaukana turvotus on tarkannäön keskipisteestä (Diabeettinen retinopatia. Käypä hoito –suositus 2024).

2.2.2 Seulontakuvaus silmänpohjakuvilla

Seulonnan tarkoituksena on löytää hoidettava tauti aikaisemmin, kuin jos se todettaisiin pelkkien oireiden ilmaantuessa. Seulonnassa käyneet jaetaan tutkimuksen jälkeen kahden joukkoon: hoidettaviin ja ei hoidettaviin. Seulonnan tarkoituksena on muun muassa parantaa elämänlaatua. (Hakama & Malila 2008.) Diabeettisessa retinopatiassa elämänlaadun parantaminen tarkoittaa mahdollisimman hyvän näkökyvyn säilyttämistä ja ylläpitämistä. Oikein ajoitettu hoito on tehokasta ja edesauttaa näkökyvyn säilymistä. (IDF 2015.) Alkuvaiheen oireettomuuden vuoksi diabeettisen retinopatian säännöllinen seulontakuvaus on perusteltua. Retinopatian eteneminen voidaan tällöin havaita ajoissa ja sitä voidaan hidastaa parantamalla kaikin keinoin diabeteksen hoitotasapainoa tai puuttua etenemiseen laserhoidolla tai lasiaiseen pistettävällä lääkeinjektioilla. (Seppänen 2021.) Diabeettista retinopatiaa ei voi seurata laboratoriotutkimuksella, vaan valokuvilla tai kliinisellä tutkimuksella (Tarnanen ym. 2024).

Jotta seulonta on laadukasta, sen tulisi täyttää vähintään seuraavat kriteerit: Olla kattavaa (yli 90 % koko diabetes populaatiosta osallistuu), oikea-aikaista (yli 90 % osallistuu kuvaukseen kolmen kuukauden aikaikkunassa suositellusta kuvauksen ajankohdasta) (Summanen & von Wendt 2014.) sekä herkkyuden olla yli 80 % sekä tarkkuus yli 94 % (Diabeettinen retinopatia. Käypä hoito -suositus 2014). Silmänpohjakuvauksessa on monia etuja. Se on luotettavaa, ja sen herkkyys on parempi kuin oftalmoskoopin antama kapea näkyvyys. Kuvista erikoislääkäreiden konsultointi on helppoa. Seulontakuvaus kattavuus paranee, kun silmänpohjakamera on liikuteltavissa. (von Wendt 2014.) Digitaalisen silmänpohjakameran käyttö ei vaadi laajaa koulutusta, jolloin tehtävään voidaan opettaa laaja-alaisesti eri terveydenhuollon ammattilaisia, esimerkiksi sairaanhoitajia tai optikoita (IDF 2015). Kuvauslaitteita on myös saatavilla optikkoliikkeissä ja silmäsairaaloissa, jolloin suuria investointeja ei tarvita (Gawęcki 2021:231).

Silmänpohjakuvauksen kultaisen standardin mukaan riittävä alue silmänpohjasta tulee kuvatuksi, kun kuvia otetaan seitsemän kentän verran 30 asteen alueen kattavilla kuvilla. Seulonnassa riittävä kuvausalue on määritelty kattamaan vähintään kaksi kuvaa otettuna 45 asteen kuvakentällä. Näin kuvattuna toinen kuvista otetaan tarkannäön keskipisteeseen tai hiukan temporaalisuuntaan päin ja toinen keskitetään papillaan tai hieman nasaalisesti. Mitä laajemmalla kuvakentällä kuvat pystytään ottamaan, sen vähemmän jää silmänpohjaa kuvaamatta ja mahdollisia retinopatian muutoksia kuvien ulkopuolelle. (Diabeettinen retinopatia. Käypä hoito –suositus 2024.)

Diabetesta sairastavat lapset kuvataan ensimmäisen kerran 10–11-vuoden ikäisenä. Kuvausväli riippuu diabetestyyppistä ja diabeettisen retinopatian tilanteesta. Mikäli muutoksia ei ole, tyypin 1 diabetesta sairastavat kuvataan kahden vuoden välein. Mikäli hoitotasapaino on pitkään hyvä, voi kuvausväli olla jopa 3 vuotta. Varhaisimmista muutoksista kuvausväli lyhenee yhteen vuoteen ja tarvittaessa tiheämmin, mikäli silmänpohjien tilannetta halutaan seurata tarkemmin. (Seppänen 2021; Diabeettinen retinopatia. Käypä hoito –suositus 2024.) Tyypin 2 diabetesta sairastavat kuuluisi kuvata piakkoin diagnoosin jälkeen, sillä retinopatiamuutoksia saattaa olla jo diagnoosihetkellä. On tyyppillistä, että tyypin 2 diabetes diagnosoidaan muutaman vuoden viiveellä, sillä oireet ovat lievemmat kuin tyypissä 1 (WHO 2024). Alle 25-vuotiaan tyypin 2 diabetesta sairastavan seulontavälit ovat kuin tyypin 1 diabetesta sairastavalla. Yli 25-vuotiaalla tyypin 2 diabetesta sairastavalla seurantaväli on kolme vuotta, mikäli muutoksia ei ole näkyvissä. Seurantaväli lyhenee kahteen vuoteen varhaisimmista muutoksista ja tarvittaessa tiheämmin. Hoitotasapainon ollessa pitkään hyvä, voi kuvausväli olla 3–5 vuoden väliltä. Raskaana olevien diabetesta sairastavien seulontakuvausta tehdään tiheämmin. Mahdollisuuksien mukaan jo etukäteen raskautta toivottaessa tai alkuraskaudessa. (Diabeettinen retinopatia. Käypä hoito –suositus 2024.) Diabeettisen retinopatian korkeampi vakavuusaste raskauden alussa on riski siihen, että raskauden aikana retinopatia etenee (Terveyskylä 2023).

Makulopatiaa epäiltäessä digitaalisten silmänpohjakuvien lisäksi hyötyä on OCT-valokerroskameran kuvista (lyhenne sanoista Optical Coherence Tomography) (IDF 2019), joka kuvaa silmänpohjan solukerrosten paksuuksia mitaten keskeisen makula-alueen turvotuksen määrän ja näyttäen sen sijainnin. (Gawęcki 2021:67–68, 70,123).

2.2.3 Seulontakuvien lausuminen

Diabeettisen retinopatian seulontakuvien lausuminen etenee muutamaa eri polkua Suomessa. Otettujen valokuvien ensimmäinen arviointi tehdään yleensä valokuvaajan

(joka on terveydenhuollon ammattilainen) toimesta, joka voi lausua kuvat itsenäisesti, mikäli kuvissa on nähtävissä korkeintaan lieviä taustaretinopatian muutoksia. Kohtalaiset ja sitä vakavammat muutokset ohjataan toiseen arviointiin silmälääkärille. Myös muut epäillyt tai tiedossa olevat silmänsairaudet lausuvat silmälääkäri. Suositeltuun kuvausväliin vaikuttavat diabetestyyppin ja retinopatian asteen lisäksi diabeteksen hoitotasapaino, raskaus sekä mahdolliset silmien hoidot ja lisätutkimukset. (Diabeettinen retinopatia. Käypä hoito –suositus 2024.) Myös muutosten sijainnilla on merkitystä, lausuuko kuvat valokuvaaja vai silmälääkäri ja mikä kuvausväli asetetaan seurannalle. Jos muutokset sijaitsevat tarkannäön tai papillan läheisyydessä, arvion tekee silmälääkäri ja kuvausväli on tällöin usein lyhyempi. (Gawęcki 2021:47).

2.2.4 Diagnostisen testin herkkyys ja tarkkuus

Seulontatutkimuksen herkkyys eli sensitiivisyys (englanniksi sensitivity) kuvastaa sitä, kuinka testi löytää positiiviset tautitapaukset positiivisten joukosta. Mikäli testi on hyvin herkkä, silloin myös väärät tautiepäilyt eli väärin positiivisten osuus kasvaa. (Suomen Lääkäriliitto 2021.) Väärät positiiviset kasvattavat turhien jatkolähetteden määrää ja samalla kustannukset nousevat (Wang ym. 2024). Seulontatestillä tulisi olla 100 % herkkyys ja myös tarkkuus tulisi olla hyvä, jotta seulonta olisi ihanteellista (Pajunpää 1999: 27).

Tarkkuus eli spesifisyys (englanniksi specificity) tarkoittaa sitä, kuinka testi tunnistaa negatiiviset tautitapaukset negatiivisten joukosta. Testin korkea tarkkuus aiheuttaa väärä negatiivisia tuloksia (Suomen Lääkäriliitto 2021.), jotka voivat olla potentiaalisesti vaarallisia ja näkökykyä uhkaavia. Korkeampi herkkyys tulisikin painottaa tarkkuuden edelle, vaikka tämä nostaa kustannuksia tarkkuuden laskun myötä (Wang ym. 2024.) Toisaalta korkea tarkkuus on hyödyllinen tilanteissa, joissa rahalliset resurssit ovat tiukat ja turhia jatkolähetteidä pitää pyrkiä välttämään (Curran ym. 2023). Tasapaino testin herkkyyden ja tarkkuuden välillä on arvovalinta, ja on hyvä pitää mielessä väärän seulontatuloksen aiheuttamat haitat potilaalle ja yhteiskunnalle. (Suomen Lääkäriliitto 2021.)

Edellisen, vuodelta 2014 olevan diabeettisen retinopatian Käypä hoito -suosituksen mukaan seulonta on laadukasta, kun sen herkkyys ylittää 80 % ja tarkkuus ylittää 94 % (Diabeettinen retinopatia. Käypä hoito -suositus 2014). Määritelmä kuitenkin poikkeaa maasta toiseen. Esimerkiksi Englannissa hyvänä seulonnan raja-arvona on pidetty vä-

hintään herkkyyttä 85 % ja tarkkuutta 80 % (Meredith ym. 2023). Tähän kirjallisuuskatsaukseen määriteltiin selkeyden vuoksi molempien raja-arvoksi 80 %, jolloin tämän ylittävät arvot ajateltiin riittävinä ja alittavat huonoina suorituksina.

2.3 Tekoäly

Tekoälyn määritelmä on vaikeaa tehdä yksiselitteisesti. Yleisesti tekoälyllä tarkoitetaan älykkyyttä. Tarkemmin ihmisenkaltaista älykkyyttä koneessa tai ohjelmistossa ja tietokoneen kykyä suoriutua samoista tehtävistä ihmiseen verrattuna. Voidaan myös määritellä, että tekoäly on ”tarkoituksenmukaista toimintaa tietokoneohjelmalla” (Toivonen 2023: Kysymys 0). Tekoälyltä yleisesti vaaditut ominaisuudet ovat autonomia (tehtävien suoritus ilman ihmisen jatkuvaa puuttumista) sekä adaptiivisuus (suorituskyvyn parantaminen kokemuksen mukaan) (Elements of AI Luku 1). Sitä mukaa, kun tekniikan kehitys on mennyt eteenpäin, on myös tekoälyltä vaaditut ominaisuudet kasvaneet. Tekoälyn on mahdollista tehdä elämästä ja työelämästä sujuvampaa, kun tekoäly voi tehdä osan rutiinitehtävistä. Tällä hetkellä tekoälyn vahvuutena on tarkasti rajatut tehtävät, jotka toistuvat ja jotka ovat samankaltaisia. Osaaminen koostuu ihmisen tekemästä ohjelmoinnista ja kouluttamisesta päätöksentekoon. Tekoälyn käyttämisen haasteena on muun muassa vastuukysymykset, kuka on vastuussa, jos kone tekee kohtalokkaan virheen? (Kolari & Kallio 2023: 14–17.)

Tällä hetkellä tekoäly on ns. heikkoa tekoälyä, jolloin se pystyy suoriutumaan vain sille ohjelmoiduista tehtävistä, toki monesti tehokkaammin kuin ihminen. Vahvaan tekoälyyn vaadittavat itsenäisen päätöksentekokyvyn ja toiminnan sekä laajan ymmärryksen näkökulmat vielä puuttuvat. (Kolari & Kallio 2023: 22.) Toisin sanoen tekoäly toimii parhaiten yhteistyössä ihmisen kanssa. Tekoäly kykenee käsittelemään yhtä aikaa massoitain tietoa ja kykenee siten tehtäviin, johon ihmisellä menisi vuosikymmeniä. Ihminen puolestaan pystyy ongelmatilanteessa näkemään kokonaisuuden, käyttämään maalaisjärkeä ja priorisoimaan oleelliset asiat epäolennaisista. (Toivonen 2023: Kysymys 2.) Tekoäly on ohjelmointia, jossa jokaiselle tarpeelle täytyy kehittää oma ohjelma tai sovellus. Tällöin tekoäly ei ole yksi yhtenäinen määrite, joka osaa kaiken, vaan tekoäly on joukko sille opetettuja erillisiä ohjelmia. (Toivonen 2023: Kysymys 40.)

2.3.1 Tekoäly terveydenhuollossa

Tekoälyä käytetään terveydenhoidossa laajalla spektrillä yksityishenkilöistä lääketieteen ja teollisuuteen. Tekoälyä voidaan käyttää muun muassa kehittäessä uusia lääk-

keitä, diagnoosien teossa sekä yksilön hyvinvointiin. Ennakoivaan hoitoon panostaminen vähentää tulevia hoitokuluja. Yhdysvalloissa ja Englannissa on tutkittu, että tekoäly tunnistaa syöpiä kuvista jo lääkäreitä tarkemmin. Leikkausrobotit puolestaan auttavat suoriutumaan leikkauksista pienemmin vaurioin kudoksiin. (Kolari & Kallio 2023: 92–95.)

Varsinkin kuvantamisessa tekoälyn hyödynnettävyys on tunnistettu. Esimerkiksi mammografiakuvia on syötetty tekoälyn oppimisdataan massoittain, jolloin tekoäly alkaa löytää ja etsiä kuvista tiettyjä rintasyöpään viittavia muutoksia. Tekoäly voi parhaimmillaan löytää kuvista varhaiset riskit tulevan rintasyövän kehittymiselle jo vuosia ennen kuin syöpäsoluja alkaa näkyä ihmissilmälle. Radiologi saattaa arvioida uransa aikana tuhansia mammografiakuvia, mutta tekoälyllä ei ole tällaista rajaa. Mitä enemmän tekoälylle syöttää erilaisia kuvia ja kertoo, onko niissä syöpämuutoksia vai ei, sen tarkemmaksi tekoäly diagnooseissaan oppii ja sen tekemät ennusteet täsmentyvät. Tekoäly on riippuvainen sille syötetyn datan oikeellisuudesta ja monipuolisuudesta. Jos sille opettaa vaikka väärän positiivisen, sen tarkkuus huononee. Jos tekoälylle opettaa vain helpot ja selvät tapaukset, se ei osaa ennustaa vaikeita kuvia. (Gospic & Passmore 2022.)

2.3.2 Tekoäly diabeettisen retinopatian seulontakuvauksessa

On olemassa useita tekoälyyn pohjautuvia sovelluksia, jotka arvioivat ja luokittelevat silmämepohjakuvia. Niiden tarkoituksena on havaita kuvat, joissa ei ole muutoksia tai vain maksimissaan lieviä muutoksia, jotka eivät uhkaa näkökykyä. Sovellukset erottelevat näköä uhkaamattomat kuvat pois ja jatkoarvioon silmälääkärin arvioitaviksi menevät vakavammat muutokset kuvissa. Olemassa olevia sovelluksia on muun muassa IDx-DR, Bosch DR, EyeArt, Google, Retmarker DR, Retinalyze ja SERI-NUS. Yhteistä sovelluksille on ollut, että niiden herkkyys tunnistaa sairas silmä on riittävän hyvä eli yli 90 %, mutta haasteita on tarkkuudessa, jolloin terveitä silmämepohjia pääsee ikään kuin seulan läpi jatkoarvioon ihmisen suorittamaan analyysiin. (Gawęcki 2021:234.) Eräässä tuoreessa 34 tutkimusartikkelia käsittelevässä kirjallisuuskatsauksessa meta-analyysin yhteenvetona oli tekoälyn herkkyys 94 % ja tarkkuus 89 %. Eli tekoäly havaitsi 94 % oikein diabeettisen retinopatian esiintymisen sekä havaitsi oikein 89 % diabeettisen retinopatian puuttumisen (Sanil ym. 2024.)

Tekoälyn käyttöä diabeettisen retinopatian tunnistamisessa on tutkittu ja tiedossa on, että tekoälypohjaisia sovelluksia on käytössä esimerkiksi Yhdysvalloissa. IDx-DR (uusi nimi LumineticsCore) on tarkoitettu aikuisille diabetesta sairastaville potilaille, joilla ei

ole tiedossa muita silmäsairauksia tai pitkälle edennyttä diabeettista retinopatiaa. Näkyvyys silmänpohjaan tulee myös olla hyvä, joten esimerkiksi kaihi tai lasiaisen samentumat estävät tekoälyalgoritmin käytön. Myös raskaana olevien silmänpohjakuvia ei tule arvioida tekoälyn avulla nopean progressiovaaran vuoksi. IDx-DR on suunniteltu tunnistamaan lievää suuremman diabeettisen retinopatian muutokset, sekä myös makulan turvotuksen. Sovelluksesta on ajateltu olevan hyötyä varsinkin perusterveydenhuollossa, jossa ei välttämättä ole kokemusta silmistä ja kuvien lausumisesta. Yhdysvaltain elintarvike- ja lääkevirasto FDA:n mukaan sovellus suoriutui 900 potilaan testissä 87,4 % herkkyydellä ja 89,5 % tarkkuudella. (FDA 2018.) Myös Suomessa on otettu käyttöön tekoälyä käyttävä silmänpohjakamera vuodenvaihteessa 23–24 Oulun yliopistollisessa sairaalassa (Suistola 2024).

Aiemmat kirjallisuuskatsaukset tekoälyn käytöstä diabeettisen retinopatian luokittelussa ovat keskittyneet tulosten tulkintaan lähinnä määrällisestä näkökulmasta, painottuen tekoälyn kuvatulkinnan herkkyyteen ja tarkkuuteen. Tässä opinnäytetyössä on tarkoitus tuoda mukaan myös laadullisin menetelmin, aineistolähtöisen sisällönanalyysin kautta, millaisia hyötyjä ja haittoja tekoälyn käytössä voi olla diabeettisen retinopatian seulontakuvauksessa ja mitä siinä tulisi ottaa huomioon. Tarkoituksena on syventää ja monipuolistaa tietoa tekoälyn hyödyntämisestä. Tekoäly kehittyy koko ajan ja sitä tutkitaan ympäri maailmaa jatkuvasti, ja näin ollen on perusteltua tehdä uusi kirjallisuuskatsaus suomeksi painottuen uusimpiin tutkimuksiin.

3 Kirjallisuuskatsauksen tarkoitus ja tavoitteet

Opinnäytetyön tarkoituksena on selvittää scoping -kirjallisuuskatsauksen keinoin, miten tekoälyä pystytään hyödyntämään diabeettisen retinopatian vaikeusasteen luokittelussa. Opinnäytetyön tavoitteena on tuottaa tietoa, voiko tekoälyä käyttää kuvien luokittelussa korvaamaan osittain silmälääkärin, tai muun lausujan tekemät lausunnot. Tavoitteena on myös tuottaa tietoa, onko tekoäly tarpeeksi luotettava ja tarkka analyysiseisään, sekä mitä hyötyä tai haittaa tekoälyn käytöstä on ilmennyt. Opinnäytetyön tulosta voi mahdollisesti hyödyntää koko Suomen hyvinvointialueilla sekä julkisella että yksityisellä sektorilla, kun suunnitellaan seuraavia hankintoja silmänpohjien kuvauslaitteistoon tai sovelluksiin. Tekoäly saattaa tehostaa diabeettisen retinopatian seulontakuvausta, kun kuvausmääriä voidaan mahdollisesti lisätä, jos terveydenhuollon ammattilaisten kapasiteettia vapautuu lausumisien automatisoinnin vuoksi. Opinnäytetyössä keskitytään löytämään vastaukset seuraaviin kahteen tutkimuskysymykseen:

1: Kuinka tarkkaa ja luotettavaa luokittelua tekoäly pystyy tekemään diabeettisen retinopatian vaikeusasteen määrittelyssä?

2: Mitä hyötyä tai haittaa on ilmennyt tekoälyn käytöstä diabeettisen retinopatian seurantakuvauksessa?

4 Kirjallisuuskatsauksen toteutus

Kirjallisuuskatsaus on tutkimus, jossa tutkitaan olemassa olevia alkuperäisiä tutkimuksia. Tarkoituksena on yhdistellä ja arvioida tietoa ja lopulta analyysin kautta tiivistää ja tuottaa uutta tietoa. Yhteistä kaikille kirjallisuuskatsauksille on niiden systemaattisuus, kurinalaisuus ja läpinäkyvyys. Kurinalaisuus näkyy siinä, että kirjallisuuskatsauksen aineisto löydetään ennakkoon määriteltujen sääntöjen ja rajoitusten avulla johdonmukaisesti. Läpinäkyvyys mahdollistuu sillä, että kaikki tehdyt valinnat perustellaan ja avataan työn kirjallisessa osassa. (Vilkkä 2023: 11–13.)

Tämän opinnäytetyön tavoitteen selvittämiseksi soveltuu hyvin kuvailevan kirjallisuuskatsauksen alalajina scoping –katsaus, jossa yksi tutkija tekee työn alusta loppuun. Scoping –katsaus kartoittaa laadullisin menetelmin olemassa olevaa näyttöä vetämällä yhteen tutkimustulokset mahdollisesti synteisiin. Scoping –katsaus on myös toteutettavissa opinnäytetyöhön varatulla melko rajatulla aikaikkunalla. (Pitkäjärvi 2024.) Tässä opinnäytetyössä tiedonhaku pyritään tekemään mahdollisimman systemoidusti ja löydetyn aineiston laadut analysoidaan, jotta kirjallisuuskatsaus olisi kokonaisuutena luotettavampi.

Kirjallisuuskatsauksessa on käytännössä kolme osaa: Tiedon keräämisen vaiheessa on minimoitava harha, mahdollisimman kattavalla ja puolueettomalla alkuperäistutkimusten otannalla. Toisessa vaiheessa on tärkeää arvioida löydettyjen tutkimusten laadut. Viimeiseksi, kun tutkimustulokset yhdistetään, se pyritään tekemään mahdollisimman yksinkertaisesti, jolloin tietoa voidaan hyödyntää tehokkaasti. (Mäkelä & Varonen & Teperi 1996.) Viimeisen vaiheen voi vielä pilkkoa kahdeksi, jolloin tutkimustulosten yhdistäminen tehdään synteisin ja analyysin kautta (Vilkkä 2023: 17). Kirjallisuuskatsauksessa systemoitu työtapo tarkoittaa sitä, että koko prosessi suunnitellaan tarkasti ja sen eteneminen kuvataan läpi kohta kohdalta. Opinnäytetyö on pystyttävä toistamaan uudelleen seuraamalla prosessi läpi. (Mäkelä ym. 1996; Vilkkä 2023: 42.)

4.1 Tiedonhaun kuvaus

4.1.1 Tietokannat

Tietokantavalinta tehdään sen perusteella mistä on oletus löytää kattavin tulos. Suosituksena on myös valita vähintään kaksi tietokantaa ja pyrkiä valitsemaan ominaisuuksiltaan erilaiset tietokannat. Tietokannoissa on eroja niiden painotuksen suhteen, eli millaisia tutkimuksia niissä julkaistaan pääsääntöisesti. On hyvä pitää mielessä myös opinnäytetyön rajalliset resurssit käytettävän ajan suhteen, kun tekee tietokantavalinnat. (Vilkka 2023: 63, 67.)

Metropolia Ammattikorkeakoulun käytössä oleviin sosiaali- ja terveystieteiden keskeisimpiin tietokantoihin (Metropolia Libguides 2024b) tutustuttiin etukäteen tekemällä harjoitushakua yksinkertaisesti hakusanoilla “artificial intelligence” AND “diabetic retinopathy” sekä “diabeettinen retinopatia” AND tekoäly. Harjoitushakujen tulosten osuvuuden ja kattavuuden perusteella tähän kirjallisuuskatsaukseen valitut neljä tietokantaa ovat Cinahl, Taylor & Francis Online, Sage Journals sekä Medic.

4.1.2 Hakusanat

Tiedonhaun tukena käytettiin apuna Metropolia Ammattikorkeakoulun kirjaston informaatiikkaa. Henkilökohtaisessa tiedonhaun opastuksessa käytiin läpi tutkimuskysymykset, relevantit hakusanat ja sopivimmat tietokannat. Informaation käyttö opinnäytetyössä parantaa sen laatua sekä luotettavuutta (Siltanen ym. 2023:14).

Hakusanoissa täytyy tunnistaa arkikielessä käytetyt sanat sekä akateemiset termit ja oikea käsitteistö. Hakutermin muodostamiseksi kannattaa käyttää synonyymeja, jotta löytää mahdollisimman kattavasti olennaiset ja sopivat artikkelit. (Vilkka 2023: 56.) Hakusanojen löytämiseksi opinnäytetyön tutkimuskysymyksistä tunnistettiin avainsanoja. Näistä avainsanoista etsittiin synonyymeja käyttäen apuna suomenkielistä YSA (Yleinen suomalainen asiasanasto) sekä englanninkielistä MeSH (medical subject headings) sivustoja. Hakuja suoritettiin tietokannoista vielä jokaisella sanalla erikseen ja määriteltiin lisää hakusanoja löydettyjen potentiaalisten tutkimusartikkeleiden asiasanoista. Taulukkoon 1 on koottu asiasanoja ja niiden synonyymeja suomeksi ja englanniksi.

Taulukko 1. Tutkimuskysymyksistä tunnistetut asiasanat ja niiden synonyymeja PICossa.

P = population, kohdejoukko	I = interest, mielenkiinto	Co = konteksti, ympäristö
Diabeettinen retinopatia	Tekoäly Keinoäly Koneoppiminen robotiikka	Seulontakuvaus Seulontatarkastus Joukkotarkastus
Diabetic retinopathy Diabetes mellitus retinopathy DMR DR	Artificial intelligence AI Machine intelligence Machine learning Deep learning Computational intelligence Computer reasoning Robotics Neural network	Screening Screening imaging Medical imaging Diagnostic imaging Classification

Hakulausekkeen muodostamisessa käytettiin apuna PICO -menetelmää, jossa tutkimuskysymykset paloitellaan sopiviksi yksittäisiksi hakusanoiksi. Hakulausekkeessa käytetään Boolean operaattoreita AND ja OR laajentamaan hakutuloksia sekä lainausmerkkejä sitomaan sanoja yhteen sekä myös sulkumerkkejä, jotta osumat saadaan mahdollisimman relevanteiksi tutkimuskysymyksiä ajatellen (Metropolia Libguides 2024a). Hakulausekkeen muodostaminen AND ja OR –sanoilla voi tuoda ristiriitaisiakin ja tutkimuskysymyksiin sopimattomia hakutuloksia (Vilka 2023: 58), mutta sen vuoksi tutkimusartikkelien sopivuus tähän kirjallisuuskatsaukseen arvioidaan seuraavan luvun sisäänotto - ja poissulkukriteerien mukaisesti.

Koehakuja suoritettiin Cinahl tietokannassa 14.4.2024 tutkimuskysymyksistä tunnistetuilla englanninkielisillä asiasanoilla ja niiden synonyymeilla. Koehakujen perusteella hakulausekkeeksi muodostui PICosta pelkästään kohdejoukko ja mielenkiinto: (**“diabetic retinopathy” OR DR**) **AND** (**“artificial intelligence” OR AI OR “machine learning” OR “deep learning” OR robotics OR “neural network”**). Haku tehtiin 1. kesäkuuta 2024 Cinahl Complete, Taylor & Francis Online sekä Sage Journals tietokannoista. Medic tietokantaan tehtiin haku samana päivänä hakulausekkeella: **“diabeettinen retinopatia” AND (tekoäly OR keinoäly OR koneoppiminen OR robotiikka)**. Tietokannoissa on eroja esimerkiksi siinä, mihin tekstinosiin haku kohdistuu (Vilka 2023: 65). Taulukossa 2 on kerrottu, millaisia hakurajauksia on käytetty tietokannoissa.

Taulukko 2. Valitut tietokannat sekä niissä käytetyt hakurajaukset. Taulukosta näkee, kuinka hakutulos pienenee sopivilla hakurajauksilla.

Tietokanta, (tulos ilman rajoituksia)	Hakurajaukset	Tulokset
Cinahl Complete (n=621)	Peer review, research article, 2020–2024, englanti	n=163
Taylor & Francis Online (n= 97 269)	Tiivistelmätason haku, 2020–2024	n=85
Sage Journals (n=56 744)	Research article, tiivistelmätason haku, 2020–2024	n=33
Medic (n=351)	Alkuperäistutkimus, 2020–2024	n=3

Valituilla hakurajauksilla pyrittiin siihen, että tietokannoista löytyvät tutkimusartikkelit kohdistuvat tarkemmin opinnäytetyön aiheeseen ja samalla niiden osuvuus tutkimuskysymyksiin paranee ja myös läpikäytävien tutkimusartikkeleiden määrä pienenee.

4.1.3 Sisäänotto- ja poissulkukriteerit

Tiedonhaussa jokainen valinta täytyy olla perusteltu etukäteen määriteltyjen sisäänotto- sekä poissulkukriteerien mukaisesti. Tutkimusartikkelit seulotaan ensin otsikkotasolla, sitten tiivistelmätasolla ja lopulta koko tekstin tasolla. (Vilka 2023: 69.) Seulontavaihe on kriittinen, varsinkin jos työ suoritetaan yksin, sillä siinä käytetään paljon tutkijakoh- taista harkintaa. (Vilka 2023: 74.) Valintakriteerejä voivat olla esimerkiksi tutkimuksen saatavuus (maksun takana, tai rajattu Metropolia Ammattikorkeakoulun pääsyn ulko- puolelle), konteksti ja soveltuvuus peilaten tutkimuskysymyksiin. Olennaista ei ole löytää tiettyä määrää tutkimuksia vaan niiden sisällön vastaavuus omiin tutkimuskysymyk- siin. Määritellyt hakurajaukset tulee olla tarkoituksenmukaisia opinnäytetyön tavoitteen ja tarkoituksen selvittämiseksi. Sama tutkimus voi olla esillä monessa tietokannassa, joten duplikaatit voi poistaa. (Vilka 2023: 69–71.) Etsittäessä tutkimuksia kirjallisuus- katsaukseen, on tarpeen tehdä alustavaa laadunarviota tutkimuksille. Mikäli kirjallisuus- katsaukseen päästää läpi heikkolaatuisia tutkimuksia, saattaa kirjallisuuskatsauksen pätevyys heikentyä. Tutkija voi määritellä itse, mutta perustellen, mitkä alkuperäistutki- musten puutteet ovat kriittisiä, jotta poissulkukriteeri täyttyy (Vilka 2023: 92–93).

Taulukko 3 mukaisten mukaanotto- ja poissulkukriteerien mukaisesti arvioitiin tietokan- noista löydettyjen tutkimusartikkelien soveltuvuus kirjallisuuskatsaukseen. Määriteltyi- hin mukaanottokriteereihin kuuluu tutkimusartikkeleiden julkaisu vuosi 2020–2024, sillä

tekoäly kehittyä vauhdilla, ja on tämän vuoksi perusteltua ottaa mukaan mahdollisimman tuoreita tutkimuksia (Kolari & Kallio 2023: 21). Julkaisukielen tulee olla suomi tai englanti opinnäytetyön tekijän osaamisen mukaisesti. Vertaisarvioinnin lisääminen mukaanottokriteereihin lisää alkuperäistutkimuksen luotettavuutta, ja näin ollen tekeillä olevan kirjallisuuskatsauksen luotettavuutta (Vilka 2023: 70). Alkuperäisjulkaisun huomiointi poistaa muun muassa muut tehdyt kirjallisuuskatsaukset pois. Tutkimusartikkelien tulee vastata etukäteen määriteltyihin tutkimuskysymyksiin, jotta ne täyttävät mukaanottokriteerit (Vilka 2023: 69). Opinnäytetyöhön varattavissa olevan ajan perusteella mukaan otetaan tutkimusartikkelit, jotka ovat verkosta löydettävissä ilman kaukotilausta. Kaikki löydetty tutkimusartikkelit olivat kuitenkin verkosta saatavilla, eikä kaukotilausta olisi edes tarvittu. Opinnäytetyöllä ei myöskään ole sponsoria ja budjettia, joten artikkeleiden on löydettävä verkosta maksutta. Maksun takana olleita tutkimusartikkeleita etsittiin useammasta tietokannasta, sekä hakemalla tutkimusartikkelin nimellä hakukonepalvelu Googlesta ja Google Scholarista. Vasta, kun varmistui ettei tutkimusartikkeli ole vapaasti saatavilla, se jätettiin pois poissulkukriteereiden mukaisesti.

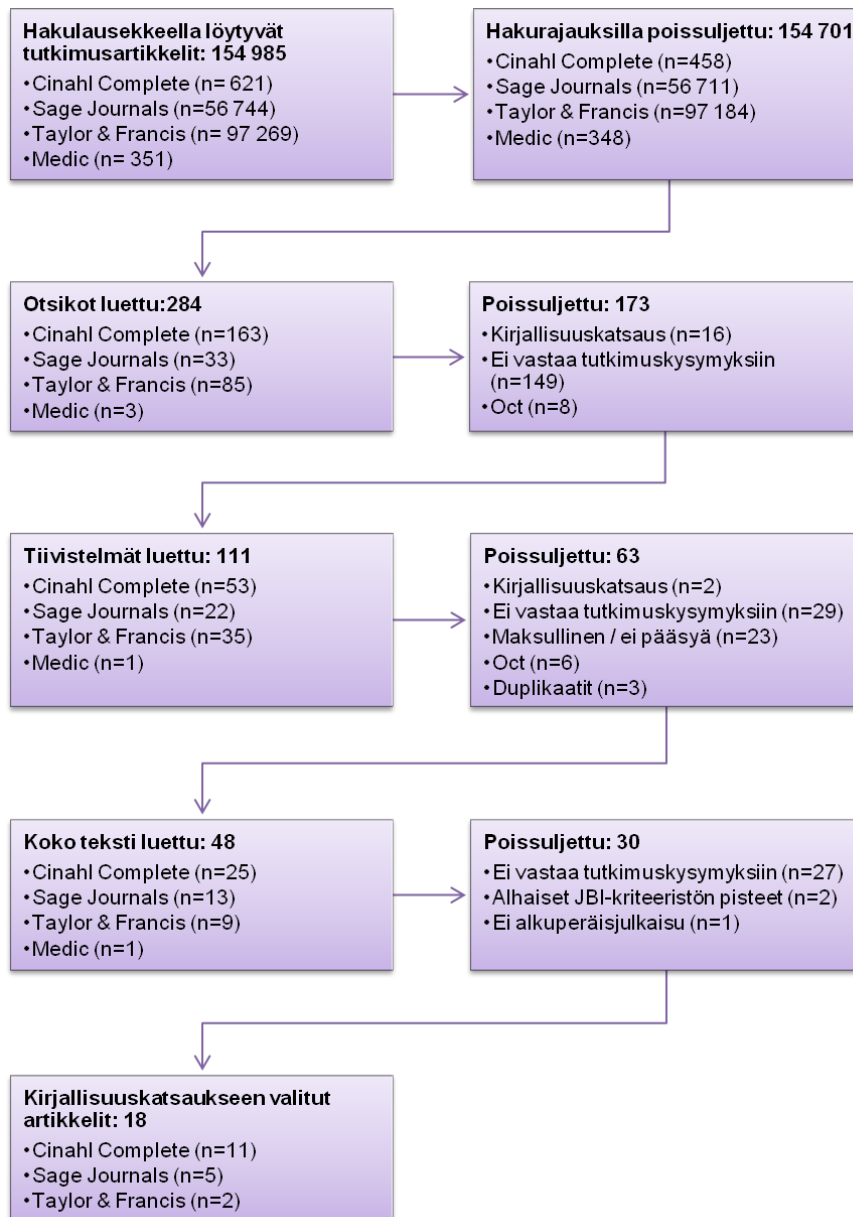
Taulukko 3. Tietokantahaussa tutkimusartikkeleiden valinnassa käytetyt mukaanotto- ja poissulkukriteerit.

Mukaanottokriteerit tietokantahaussa	Poissulkukriteerit tietokantahaussa
<ul style="list-style-type: none"> • Julkaistu vuonna 2020–2024 • Julkaisukieli suomi ja englanti • Vertaisarvioitu tieteellinen artikkeli • Alkuperäisjulkaisu • Vastaa tutkimuskysymyksiin • Artikkelit löydettävissä kokonaan ja full text saatavilla • JBI- laadunarvioinnissa vähintään puolet pisteistä 	<ul style="list-style-type: none"> • Julkaistu ennen vuotta 2020 • Julkaisukieli muu, kuin suomi tai englanti • Ei vertaisarvioitu • Muu kuin alkuperäisjulkaisu • Ei vastaa tutkimuskysymyksiin • Artikkelit ei löydy kokonaisuudessaan luettavaksi tai se on maksullinen • JBI- laadunarvioinnissa ei täyty puolet pisteistä

Haettaessa tutkimusartikkeleita tietokannoista, voi hakulausekkeella tulla ristiriitaisia tai omiin tutkimuskysymyksiin vastaamattomia tuloksia. Jokainen osuma tietokannoista tarkasteltiin läpi mukaanotto- ja poissulkukriteerien mukaisesti.

4.1.4 PRISMA –kaavio

Tiedonhaku tehostetaan järjestelmällisellä haulla. Hyödyntämällä PRISMA–kaaviota, hakua havainnollistetaan, sekä myös kirjallisuuskatsauksen toistettavuus ja samalla sen luotettavuus paranee. (Vilkkä 2023: 67.) 1. kesäkuuta 2024 tehty tiedonhaku on kuvattu PRISMA-kaavioon kuvioon 2, jossa on eriteltynä tietokannoista löydetty tutkimusartikkelit kappalemäärittäin ja pois sulkemiset kriteereittäin. N-määrät on täytetty toteutuneen tietokantahaun mukaisesti. Haun jälkeen, kirjallisuuskatsaukseen jäi 18 tutkimusartikkelia.



Kuvio 2. Tietokannoista löydetty tutkimusartikkelit kuvattuna PRISMA flow –kaavioon mukaan-otto- ja poissulkukriteerien mukaisesti karsittuna. (mukaillen PRISMA 2024).

4.2 Aineiston laadun arviointi

Kirjallisuuskatsauksen tutkimusartikkeleiden laadunarvio on tärkeä tehdä kriittisesti. Tarkoituksena on arvioida asianmukaisuutta, mutta ei arvostella ylikriittisesti tutkimuksia. Tutkimuskysymykset ja kirjallisuuskatsauksen tarkoitus pidetään mielessä, kun tutkimusartikkeleiden vahvuuksia, heikkouksia ja soveltuvuutta ja tarkoituksenmukaisuutta omaan tutkimukseen arvioidaan. (Vilkkä 2023: 15, 92–94.)

Jokaisen tutkimusartikkelin laatu arvioitiin Joanna Briggs Instituutin (JBI) laadunarvioiden mukaisesti, jotka Hoitotyön tutkimussäätiö (Hotus) on suomentanut. Kunkin tutkimusartikkelin tutkimusmenetelmään valittiin sopivin laadunarvion lomake. Laadunarvioinneista on koontina taulukko 4, joka selventää annetut pisteet kullekin tutkimusartikkelille. Tutkimuksien laaduissa on hajontaa laadunarvioiden perusteella, mutta hyväksytyjen tutkimusten laatujen arvioitiin olevan riittäviä, jotta ne soveltuvat tähän opinäytetyön kirjallisuuskatsaukseen. Tässä opinnäytetyössä ei käsitellä kahta tutkimusartikkelia, jotka saivat laadunarvioissa alle puolet jaossa olevista pisteistä (Mokhashi ym. 2022; Wolf ym. 2020).

Taulukko 4. Tutkimusartikkeleiden laadunarvioinnit JBI-kriteerien mukaisesti. (• diagnostisen testin tarkkuustutkimus) (•• taloudellinen tutkimus) (••• kvasikokeellinen tutkimus) (•••• satunnaistettu kontrolloitu tutkimus) (K1= kysymys 1, K2=kysymys 2 jne.) Ei (E), kyllä (K), epäselvä (?), Ei sovellettavissa (-)

Tutkimus	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13	Yht.
Kubin ym. 2024 •	E	K	?	K	-	K	K	K	K	K				7/10
Musetti ym. 2024 •	E	K	K	K	-	K	K	K	K	K				8/10
Wang ym. 2024 ••	K	K	K	K	K	K	K	K	E	K	?			9/11
Dow ym. 2023 •••	K	K	K	K	K	E	?	E	E					5/9
Abramoff ym. 2023 ••••	K	K	K	K	K	K	K	K	E	K	E	K	K	11/13
Wroblewski ym. 2023 •	K	K	K	K	-	K	K	K	K	K				9/10

Curran ym. 2023 •	K	K	E	K	-	E	K	K	E	K				6/10
Teoh ym. 2023 •	E	K	?	K	-	K	K	K	K	K				7/10
Meredith ym. 2023 •	K	K	E	K	-	K	K	K	E	K				7/10
Mokhashi ym. 2022	K	K	E	?	-	E	?	K	E	E				3/10
Malerbi ym. 2022 •	K	K	E	K	-	?	K	K	K	K				7/10
Huang ym. 2022 ••	K	K	K	K	E	K	K	K	K	E	E			8/11
Malerbi ym. 2022 •	E	K	E	K	-	K	K	K	E	K				6/10
Liljenquist ym. 2021 •	K	K	E	K	-	K	K	K	K	E				7/10
Shah ym. 2021 •	K	K	E	K	-	K	K	K	K	E				7/10
Wolf ym. 2021 •	K	K	E	K	-	K	K	K	E	E				6/10
Lee ym. 2021 •	K	K	K	K	-	K	K	K	K	K				9/10
Wolf ym. 2020	K	K	E	K	E	K	E	E	K	E	E			5/11
Vaghefi ym. 2020 •	K	K	E	K	-	K	K	K	K	K				8/10
Sosale ym. 2020 •	K	K	E	K	-	K	K	K	K	K				8/10

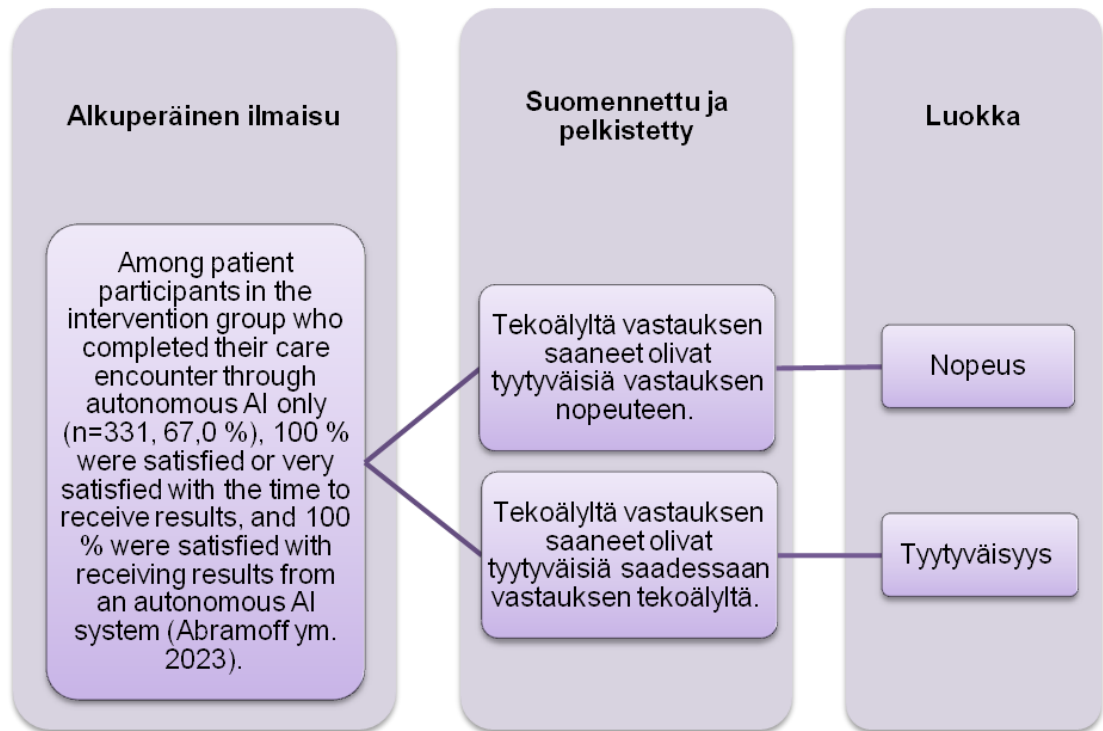
4.3 Aineiston analysointi

Kirjallisuuskatsaukseen valittiin menetelmäksi aineistolähtöinen analyysi, jossa edetään tutkimusartikkeleista löytyvien tulosten perusteella. Lopputuloksesta ei tehdä oletuksia etukäteen, vaan annetaan aineiston ohjata. (Tuomi & Sarajärvi 2018: Luku 4.2.) Sisällönanalyysin tavoitteena on saada jäsenneltyä aineisto, jotta siitä voidaan tehdä

johtopäätöksiä (Tuomi & Sarajärvi 2018: Luku 4.4). Aineistolähtöisessä aineiston analyysissä on kolme osaa: aineiston pelkistäminen, ryhmittely sekä käsitteiden luominen. Pelkistämässä tutkimusaineistosta etsitään tutkimuskysymyksiin vastaavia ilmaisuja, jotka kootaan allekkain ja näistä alkuperäisistä ilmauksista tehdään yksi tai useampia pelkistettyjä ilmauksia. Seuraavassa, ryhmittelyn vaiheessa, pelkistämisen ilmaukset ryhmitellään samankaltaisten ominaisuuksien mukaan luokiksi. Luokittelua voidaan jatkaa tarvittaessa pääluokkiin, yläluokkiin ja alaluokkiin. Ryhmittelyn jälkeen pyritään löytämään luokitteluiden avulla olennaiset seikat, joiden avulla lopulta vastataan tutkimuskysymyksiin. (Tuomi & Sarajärvi 2018: Luku 4.4.3.)

Aineistolähtöisessä sisällönanalyysissä on tarkoitus löytää tutkimuksista yhteisiä, mutta myös eriäviä lopputulemia ja pohtia niiden syitä. Myös on hyvä tunnistaa, jos joku havainto jää vahvistamatta ja siitä kaivattaisiin vielä lisää tutkimusta. Tarkoituksena on erittelyn ja vertailun keinoin tutkivalla lukemisen otteella yhdistää tutkimusartikkeleista yhteinen asiakokonaisuus ja tehdä päätelmiä. Nämä päätelmät jäsennellään yleensä visuaalisesti tai tiivistetään taulukoimalla. (Vilkkä 2023: 86–87.) Tutkimusartikkeleista tehdään muistiinpanoja tutkimuskysymykset mielessä. Lopuksi analyysin tuloksia tarkastellaan ja pohditaan vertaillen ja yhdistellen tietoa opinnäytetyön johdannon ja teoreettisen viitekehyksen kanssa (Vilkkä 2023: 80–82, 87.)

Kuviossa 3 on esitelty esimerkkinä, kuinka kirjallisuuskatsauksessa edettiin tutkimusartikkeleiden alkuperäisistä ilmaisuista suomennoksen ja pelkistykseen kautta luokitteluun. Tutkimusartikkelit luettiin läpi keskittyen kohtiin, jotka vastaavat kirjallisuuskatsauksen tutkimuskysymyksiin. Nämä tekstikohdat käännettiin ensin suomeksi. Suomentamisen jälkeen lauseita pelkistettiin ja niistä karsittiin kirjallisuuskatsaukselle epäolennaiset kohdat pois kuitenkin muuttamatta alkuperäistä sisältöä. Pelkistetyt lauseet ryhmiteltiin sen jälkeen samankaltaisten ilmausten joukoiksi. Näistä joukoista muodostettiin koko joukkiota parhaiten kuvaavia luokkia. Luokat jaettiin sen jälkeen tekoälyn käytöstä diabeettisen retinopatian seulontakuvauksessa hyötyihin ja haittoihin, joista muodostettiin omat kuvionsa 4 ja 5. Kuvioihin muodostettujen luokittelujen avulla pyrittiin lopulta vastaamaan kirjallisuuskatsauksen tutkimuskysymyksiin.



Kuvio 3. Esimerkki kirjallisuuskatsauksessa tehdystä aineiston pelkistämisestä ja ryhmittelystä aineistolähtöisessä sisällönanalyysissa. Alkuperäinen englanninkielinen ilmaisu suomennettiin, pelkistettiin ja niille luotiin kuvaavia luokkia.

Opinnäytetyössä käytetyt tutkimusartikkelit on järjestetty edempänä olevaan taulukko 5 julkaisuvuoden mukaan alkaen tuoreimmasta tutkimuksesta ja päättyen vanhimpaan. Taulukkoon on kerätty artikkelin kirjoittaja, julkaisuvuosi, julkaisumaa, tutkimuksen tavoite, tutkimustyyppi, ketä on tutkittu ja tutkimuksen päätulokset. Taulukointi oli apuna tutkimusartikkeleiden jäsentämisessä ja sisällönanalyysissa.

Taulukko 5. Tietokantahausta löydetty tutkimusartikkelit yksityiskohtineen tiivistetyssä muodossa taulukkoon koottuna.

Kirjoittaja, vuosiluku, maa	Tavoite	Tutkimus-tyyppi	Ketä tutkittu	Tutkimuksen päätulokset
Kubin ym. 2024, Suomi	21 anonyymien tekoälyalgoritmien vertailu tunnistamaan lähetettä vaativa diabeettinen retinopatia käyttäen käsikäyttöisen silmänpohjakameran kuvia. Referenssinä kahden silmälääkärin lausunnot.	Diagnostisen testin tarkkuustutkimus	156 diabetes potilasta (312 silmää)	Algoritmien suoriutumisessa suurta vaihtelua tunnistaa lähetettä vaativat kuvat: 49,4 % -92,3 %. Parhaiten suoriutuneiden viiden algoritmin keskiarvoinen herkkyys 84,3 % ja tarkkuus 91,3 %. 14 / 21 algoritmilla alle 80 % herkkyys tai tarkkuus. Kuvamäärä, jota tekoäly ei pystynyt luokittelemaan vaihteli 0–28,2 % välillä.
Musetti ym. 2024, Italia	Kahden eri tekoälyalgoritmien suoriutuminen tunnistaa ”hyvin lievää” isommat muutokset kuvista. Referenssinä kahden silmälääkärin lausunnot.	Diagnostisen testin tarkkuustutkimus	201 yli 18-vuotista diabetes potilasta	Algoritmien herkkyudet olivat vastaavat kuin ihmisellä: 100 % ja 94,3 %. Tarkkuudet alhaisemmat: 72,9 % ja 66 % (Väärät positiiviset 50 ja 64 kappaletta). Tekoälyt eivät pystyneet luokittelemaan 6,5 % ja 8 % kuvista.
Wang ym. 2024, Kiina	Tekoälyn suorituskyvyn ja pitkäaikaisen kustannustehokkuuden välistä suhdetta arvioitiin mallintamalla 1100 herkkyys/tarkkuus -paria ja vuotuisia seulontaskenaarioita. Näistä malleista tarkin määritettiin referenssiksi, johon verrattiin muut vaihtoehdot.	Taloudellinen arviointi	251 535 diabetes potilasta (865 152 kuvaa)	Referenssin herkkyys 93,3 % ja tarkkuus 87,7 %, johon verrattuna kuusi skenaariota toi kustannussäästöjä ja seitsemän kustannustehokkuutta. Herkkyys oltava yli 88,2 % (paritarkkuus 90,3 %), jotta toiminta on eniten kustannuksia säästävää ja tarkkuus yli 80,4 % (pariherkkyys 96,3 %), jotta tuottaa kustannustehokkuutta. Jos herkkyys nousee yli 96,3 % tai tarkkuus yli 90,3 %, niihin liittyvät (pari) tarkkuuden ja herkkyuden laskut heikentävät kustannustehokkuushyöty-suhdetta.
Dow ym. 2023, Yhdysvallat	Kolmen vuoden pituinen tutkimus, jossa verrataan ”ihmistyökulkua” tekoäly-työkulkuun. Tutkimuksessa seurattiin, kuinka moni kävi seurannassa 90pv sisällä lähetteen saatuaan ja miten tekoäly vaikuttaa.	Kvasikokeellinen	2243 aikuista diabetes potilasta (tekoäly). 790 aikuista diabetes potilasta (ihmisarvio)	Potilaiden osallistuminen seurantaan on kolme kertaa korkeampi tekoälyn tekemän päätöksen jälkeen (35,5 %), kuin ihmispäätöksen jälkeen (12 % ja 11,7 %). Tekoäly ei pystynyt arvioimaan 34,95 % kuvista.

Abramoff ym. 2023, Bangladesh	Hypoteesina on, että tekoäly tehostaa silmälääkärin työskentelyä. 105 vastaanottopäivää satunnaistetusti joko tekoälyn kanssa tai ilman.	Satunnais-kontrolloitu tutkimus	Kolme verkkokalvoon erikoistunutta silmälääkärää. Potilaat yli 22-vuotiaita, joilla diabetes.	Tekoäly paransi tuottavuutta 40 %: tekoälyn avulla 1.59 kohtaamista tunnissa ja ilman tekoälyä 1.14 kohtaamista tunnissa. 331 potilasta, jotka saivat tuloksen pelkästään tekoälyltä, olivat 100 % tyytyväisiä vastauksen saamiseen tekoälyltä ja myös tyytyväisiä vastauksen nopeuteen. Tekoäly ei voinut arvioida 3,42 % kuvista.
Wroblewski ym. 2023, Meksiko	Kahden tekoälyalgoritmin suoriutuminen löytää ”hyvin lievää” suuremmat muutokset kuvista. Referenssinä kahden silmälääkärin lausunnot.	Diagnostisen testin tarkkuustutkimus	248 diabetes potilasta (5–80-vuotiaita) (2130 kuvaa)	Toisen algoritmin suoriutuminen: herkkyys 94 %, tarkkuus 94 % ja pystyi arvioimaan jokaisen kuvan. Toisen algoritmin suoriutuminen: herkkyys 94 %, tarkkuus 86 %, mutta riittämätön kuvanlaatu 37 % kuvassa.
Curran ym. 2023, Bangladesh	Tekoälyn suoriutuminen lausumaan kuvista diabeettisen retinopatian muutokset lapsilta ja nuorilta. Referenssinä optikon lausunnot.	Diagnostisen testin tarkkuustutkimus	1274 diabetes potilasta (3–26-vuotiaita)	Minkä tahansa diabeettisen retinopatian asteen erottelukyvyn herkkyys tekoälyllä oli 75,5 % ja tarkkuus 91,8 %. Jatkolähetettä vaativat kuvat: herkkyys 84,2 % ja tarkkuus 98,9 %. Tekoäly oli tarkin luokittelemaan nuorimmat potilaat (keskimäärin 16.7 vuotta vastaan 22.9 vuotta) pienimmillä diagnoosivuosilla (keskimäärin 4.85 vuotta vastaan 8.50 vuotta)
Teoh ym. 2023, Singapore	Tarkoitus selvittää miten tekoäly selviää vastaan ihmiset, joilla eri osaamisen asteet luokitteluun. 200 makulakeskeistä kuvaa luokiteltiin: ei lähetettä, ei kiireinen lähete ja kiireellinen lähete. Referenssinä kolme erikoistunutta silmälääkärää.	Diagnostisen testin tarkkuustutkimus	Referenssi vastaan tekoäly, 3 erikoistuvaa silmälääkärää, 3 yleislääkärää ja 2 lääketieteen opiskelijaa	Tekoälyn antama vastaus ”ei lähetettä” herkkyys 70,1 %, ja tarkkuus 93,5 %. ”Ei kiireellinen lähete” herkkyys: 90,1 % ja tarkkuus 45,4 %. ”Kiireinen lähete” herkkyys 85,7 % ja tarkkuus 75,9 %. Tekoälyllä potentiaalia avustaa epäpäteviä tai kokemattomia kuvalausujia työssään.
Meredith ym. 2023, Englanti	Miten tekoäly selviää diabeettisen retinopatian luokittelusta englantilaisessa populaatiossa	Diagnostisen testin tarkkuustutkimus	9817 kuvasettiä (12–100-vuotiaita diabetes potilaita)	”Ei diabeettista retinopatiaa” vastaan mikä tahansa diabeettinen retinopatia herkkyys 69,7 % ja tarkkuus 92,2 %. ”Ei diabeettista retinopatiaa” tai lievä diabeettinen retinopatia vastaan lähetettä vaativa diabeettinen retinopatia herkkyys 95,4 % ja tarkkuus 92,0 %. Tekoälyltä ei jäänyt yksikään merkittävä huomaamatta. Lausui kaikki kuvat. 24 kuvassa ihmisarvio ja tekoälyn vastaus poikkesivat toisistaan merkittävästi.

Malerbi ym. 2022a, Brasilia	Käsikäyttöinen silmänpohjakamera ja tekoälyn diagnostinen tarkkuus tunnistaa diabeettinen retinopatia silmänpohjakuvista. Referenssinä yhden silmälääkärin lausunnot.	Diagnostisen testin tarkkuustutkimus	679 yli 18-vuotista 2 tyypin diabetes potilasta (3 255 kuvaa).	Tekoälyn kyky tunnistaa lievää suurempia diabeettiset retinopatiat herkkyys 97,8 % ja tarkkuus 61,4 %. (Vääriä positiivisia tekoälyllä 191/495) Kaikki tekoälyn antamat väärät negatiiviset (4 /184) olivat ihmisarviossa kohtalaisia diabeettisen retinopatian muutoksia, joista kahdella oli makulaturvotusta.
Huang ym. 2022, Kiina	Tarkoitus selvittää tekoälylähtöisen kuva-seulonnan kustannustehokkuutta vastaan "ei seulontaa" tai "silmälääkäriperusteinen seulonta". Seuranta-aika oli 35 vuoden ajan, jossa seulonnat vuosittain.	Taloudellinen arviointi	1000 uutta diabetes potilasta keski-ikänsä 44, joilla ei diabeettista retinopatiaa. (hypoteettinen tilanne)	Terveysjärjestelmän näkökulmasta tekoälyn inkrementaalinen kustannus 180,19 dollaria, silmälääkärikustannus 215,05 dollaria. (terveyshyötyjä: lisäsi terveyslaadukkaita elinvuosia vastaan ei seulontaa ollenkaan ja vastaan silmälääkärin lausunto). Tekoälyn seulonta kustannustehokkaampaa kuin silmälääkäri seulonta sekä terveysjärjestelmän, että yhteiskunnan näkökulmasta.
Malerbi ym. 2022b, Brasilia	Tarkoitus arvioida makulaturvotuksen esiintymistä 2 tyypin diabetespotilailla värivalokuvista tekoälyn avulla. Referenssinä kahden silmälääkärin lausunnot. Käytetty kamera älypuhelinperusteinen käsikäyttöinen.	Diagnostisen testin tarkkuustutkimus	366 yli 18-vuotista 2 tyypin diabetes potilasta	Tekoäly epäili turvotusta 60 silmälle, oikea positiivinen tulos makulaturvotuksen epäilyistä 47 silmää eli 78,3 %. Väärä positiivinen tulos 13 silmää eli 21,7 % (Synyvä kovat eksudaatit ilman turvotusta (8), pigmenttikertymä (4), kuvan artefakti (1))
Liljenquist ym. 2021, Yhdysvallat	Miten tekoäly selviää vastaan referenssinä kahden silmälääkärin lausunnot, kun etsitään kuvista näköä uhkaavia muutoksia tai lievää suurempia muutoksia.	Diagnostisen testin tarkkuustutkimus	893 yli 18-vuotista diabetes potilasta (1786 silmää)	Tekoälyn lievää suurempien muutosten havaitsemisen herkkyys 95,5 %, tarkkuus 85 %. Näköä uhkaavien muutosten havaitseminen herkkyys 95,1 %, tarkkuus 89 %. Suoritetun aavistus parempaa laajennettujen pupillien kanssa. Lausumismäärä nousi 87,4–97,4 % laajennustippojen avulla.
Shah ym. 2021, Espanja	Miten tekoäly suoriutuu, kun etsitään jatkolähetettä vaativia kuvia. Referenssinä kolmen silmälääkärin lausunnot.	Diagnostisen testin tarkkuustutkimus	2680 diabetes potilasta.	Tekoäly tunnisti jatkolähetettä vaativat kuvat herkkyys: 100 %, tarkkuus 81,82 % (467 väärä positiivinen). Näköä uhkaavat kuvat herkkyys 100 %, tarkkuus 94,64 % (140 väärä positiivinen). Tekoälyn tarkkuus laskee, kun diagnoosivuotia yli 10 vastaan alle 10 vuotta, tai kun potilas yli 65-vuotias vastaan alle 65-vuotias. 13,2 % kuvista jäi tekoälyltä lausumatta.

Wolf ym. 2021, Yhdysvallat	Tekoälyn diagnostinen tehokkuus lapsilla ja nuorilla diabetesta sairastavilla. Referenssinä kaksi silmälääkärää. Seurattiin myös seulontaohjeiden noudattamista.	Diagnostisen testin tarkkuustutkimus	310 diabetes potilasta. (5–21-vuotiaita)	Tekoäly pystyi luokittelemaan 97,5 % kuvista (302/310). Lievää suuremmat muutokset tekoäly tunnistui herkkyydellä 85,7 % ja tarkkuudella 79,3 %. Tekoälyn 61 väärää positiivista (kiiltelevä kalvo n=44, laatuongelmia n=17).
Lee ym. 2021, Yhdysvallat (Atlanta ja Seattle)	Seitsemän tekoälyalgoritmin suoriutumisen vertailua. Tarkoituksena tunnistaa jatkolähetettä vaativat kuvat ei lähetettä vaativista kuvista. Referenssinä optikoiden ja silmälääkäreiden lausunnot	Diagnostisen testin tarkkuustutkimus	23 724 diabetes potilasta (21–98-vuotiaita), (311 604 kuvaa)	Vaihteluväli suurta: Algoritmien herkkyydet vaihtelivat 50,98–85,90 % ja tarkkuus 60,42–83,69 %. 3/7 herkkyydeltään parempi tai vastaava kuin ihmisen tulkinta. Yhden herkkyyks oli alin (74,42 %) tunnistamaan proliferatiivinen diabeettinen retinopatia, joka tarkoittaa, että tekoäly ei 25,58 %:ssa tapauksista tunnistanut edenneitä muutoksia.
Vaghefi ym. 2020, Uusi-Seelanti	Kehittää diabeettisen retinopatian arviointityökalu, joka tunnistaa lähetettä vaativat tapaukset, myös makulopatian.	Diagnostisen testin tarkkuustutkimus	32 354 diabetes potilasta (7–10-vuotiaita) (seurantakäynteineen 63 843 potilasta)	Datasetti 1: Tekoälyn herkkyyks 94 % ja tarkkuus 63 %. Datasetti 2: Tekoälyn herkkyyks 95 % ja tarkkuus 61 %. Tekoäly arvioi 65–70 % kuvista ei vaarallisiksi eli tekoäly vähentää kuvaluokittelun työmäärää n. 65 %. Tekoälyalgoritmi lausui 94 % kuvista.
Sosale ym. 2020, Intia	Tutkia älypuhelinperustaisen tekoälysovelluksen herkkyyttä ja tarkkuutta lähetettä vaativan diabeettisen retinopatian (enemmän kuin kohdalainen) havaitsemiseen. Referenssinä kahden silmälääkärin lausunnot.	Diagnostisen testin tarkkuustutkimus	297 yli 18 v diabetes potilasta	Lähetettä vaativat kuvat tekoäly herkkyyks 98,84 % ja tarkkuus 86,73 %. Näköä uhkaavien kuvien tunnistamisen herkkyyks 100 %. Mikä tahansa diabeettisen retinopatian asteen herkkyyks 86,78 % ja tarkkuus 95,45 %. Tekoälyn väärät positiiviset n=8. 6/8 näistä oli kuvan artefakti syynä.

4.4 Tutkimusten yhteenveto

Kirjallisuuskatsaukseen valittiin 18 alkuperäistä tutkimusartikkelia, jotka saivat laadun-
arviossa yli puolet pisteistä. Tutkimukset oli julkaistu sisäänottokriteerien mukaisesti
vuosien 2020 ja 2024 välillä. Tutkimuksista 14 kappaletta käsitteli diagnostisen testin
tarkkuustutkimusta (Kubin ym. 2024; Musetti ym. 2024; Wroblewski ym. 2023; Curran
ym. 2023; Teoh ym. 2023; Meredith ym. 2023; Malerbi ym. 2022a; Malerbi ym. 2022b;
Liljenquist ym. 2021; Shah ym. 2021; Wolf ym. 2021; Lee ym. 2021; Vaghefi ym. 2020;
Sosale & Sosale & Murthy & Sengupta & Naveenam 2020). Lisäksi oli taloudellisia tut-
kimuksia (Wang ym. 2024; Huang ym. 2022), satunnaistettu kontrolloitu tutkimus (Ab-
ramoff ym. 2023) ja kvasikokeellinen tutkimus (Dow ym. 2023).

Taloudelliset tutkimukset oli tehty Kiinassa, joten niissä ei ole maantieteellistä variaa-
tiota (Wang ym. 2024; Huang ym. 2022). Loput tutkimukset ovat hajaantuneet laajem-
min: Pohjois-Amerikassa on tehty viisi tutkimusta (Dow ym. 2023; Wroblewski ym.
2023; Liljenquist ym. 2021; Wolf ym. 2021; Lee ym. 2021), Etelä-Amerikassa kaksi tut-
kimusta (Malerbi ym. 2022a; Malerbi ym. 2022b), Euroopassa neljä tutkimusta (Kubin
ym. 2024; Musetti ym. 2024; Meredith ym. 2023; Shah ym. 2021), Aasiassa neljä tutki-
musta (Abramoff ym. 2023; Curran ym. 2023; Teoh ym. 2023; Sosale ym. 2020) sekä
Uudessa-Seelannissa yksi (Vaghefi ym. 2020).

Diagnostisten testien tarkkuustutkimusten otoskoot vaihtelivat 156–32 354 potilaan vä-
lillä (Kubin ym. 2024; Vaghefi ym. 2020). Suurimmassa osassa tutkimuksia otannassa
oli vain aikuisia (Musetti ym. 2024; Dow ym. 2023; Abramoff ym. 2023; Malerbi ym.
2022a; Malerbi ym. 2022b; Liljenquist ym. 2021; Lee ym. 2021; Sosale ym. 2020) ja
harvinaisempaa oli, että otanta kohdistui pelkästään lapsiin ja nuoriin (Curran ym.
2023; Wolf ym. 2021). Sekatutkimuksia iän suhteenkin oli (Wroblewski ym. 2023; Mere-
dith ym. 2023; Vaghefi ym. 2020) ja kaikissa tutkimuksissa ei ollut ikäjakaumaa mai-
nittu lainkaan (Kubin ym. 2024; Teoh ym. 2023; Shah ym. 2021).

Yleisintä tutkimuksissa oli käsitellä yhtä tekoälyyn pohjautuvaa algoritmia (Dow ym.
2023; Abramoff ym. 2023; Curran ym. 2023; Teoh ym. 2023; Meredith ym. 2023; Ma-
lerbi ym. 2022a; Malerbi ym. 2022b; Liljenquist ym. 2021; Shah ym. 2021; Wolf ym.
2021; Vaghefi ym. 2020; Sosale ym. 2020). Myös 2–21 tekoälyalgoritmia vertailevia tut-
kimuksia oli. Niissä algoritmien suoriutumista verrattiin samaan potilasjoukkoon (Mu-
setti ym. 2024; Wroblewski ym. 2023; Lee ym. 2021; Kubin ym. 2024.) Tutkimuksissa
oli nimetty yhteensä kahdeksan eri tekoälyalgoritmia. Algoritmeja oli kuitenkin kokonai-

suudessaan vähintään 21 erilaista, koska suomalaisessa tutkimuksessa vertailtiin anonyymisti niiden suoriutumista. Näistä 21 algoritmista 14 toimittajaa oli antanut luvan julkaista nimen tutkimuksessa. (Kubin ym. 2024.) Tutkimuksissa käytettyjä erilaisia kameramalleja oli mainittu yhteensä 12, sekä eroa oli myös siinä, oliko potilaiden pupilleja laajennettu tipoilla (Abramoff ym. 2023; Wroblewski ym. 2023; Malerbi ym. 2022a; Shah ym. 2021; Sosale ym. 2020) vai ei (Dow ym. 2023; Curran ym. 2023; Wolf ym. 2021). Laajennustippojen käytöstä ei ole mainittu viidessä tutkimuksessa (Kubin ym. 2024; Teoh ym. 2023; Meredith ym. 2023; Malerbi ym. 2022b; Vaghefi ym. 2020) ja osassa oli sekä laajennettuja että laajentamattomia pupilleja (Musetti ym. 2024; Liljenquist ym. 2021; Lee ym. 2021).

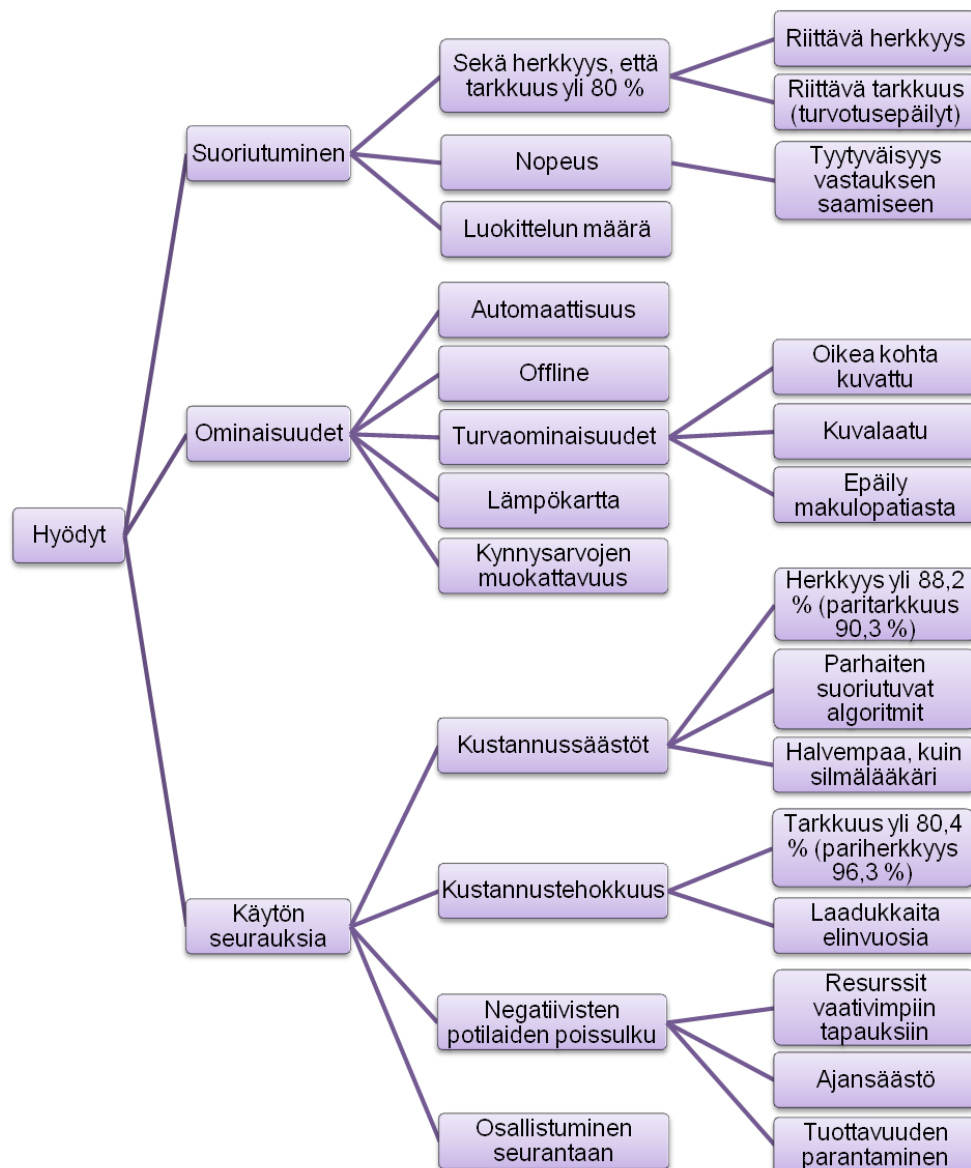
Yleisin tekoälyalgoritmin käsittelemä kuvamäärä oli kaksi kuvaa silmää kohden (Kubin ym. 2024; Musetti ym. 2024; Wang ym. 2024; Dow ym. 2023; Abramoff ym. 2023; Meredith ym. 2023; Malerbi ym. 2022a; Liljenquist ym. 2021; Shah ym. 2021; Wolf ym. 2021). Myös pelkät yhdet makulakuvat riittivät aineistoksi (Curran ym. 2023; Teoh ym. 2023; Malerbi ym. 2022b). Kuvia saattoi olla myös kolme tai enemmän silmää kohden (Wroblewski ym. 2023; Lee ym. 2021; Sosale ym. 2020). Kuvamäärä muuttui Uuden-Seelannin tutkimuksessa diabetestyyppin mukaan: 2 tyyppin diabetesta sairastavilta otettiin kaksi kuvaa ja 1 tyyppin diabetesta sairastavilta neljä kuvaa. (Vaghefi ym. 2020.)

11 tutkimuksessa oli poistettu potilasjoukosta esimerkiksi sellaisia potilaita, joilla oli tiedossa joku muu silmäsairaus, tehty silmäleikkaus tai esimerkiksi hoidettu silmänpohjaa laserilla. Myös huonolaatuisia kuvia oli jätetty pois tutkimuksien otoksesta. Syynä saattoi olla myös puuttuvat referenssitiedot tai kuvaa ei ollut otettu oikeasta kohdasta silmänpohjaa. (Abramoff ym. 2023; Curran ym. 2023; Meredith ym. 2023; Malerbi ym. 2022a; Malerbi ym. 2022b; Liljenquist ym. 2021; Shah ym. 2021; Wolf ym. 2021; Lee ym. 2021; Vaghefi ym. 2020; Sosale ym. 2020.) Tämä voi vaikuttaa tekoälyn näennäisesti parempaan suoriutumiseen, kun potilasjoukko ei ole tosielämän tilannetta vastaava, vaan jo etukäteen rajattu joukko. Potilasjoukkoa rajaamattomiakin tutkimuksia oli (Kubin ym. 2024; Wroblewski ym. 2023) tai mahdollinen rajausta ei tullut raportissa esille (Musetti ym. 2024; Dow ym. 2023; Teoh ym. 2023).

5 Tulokset

5.1 Tekoälyn käytön hyötyjä diabeettisessa retinopatiassa

Tutkimusartikkeleiden taulukoinnin jälkeen (taulukko 5) niistä eriteltiin hyödyt ja haitat omien kuvaavien luokkien alle, joista muodostettiin kuvat 4 ja 5. Kuvioon neljä on koottu tutkimusartikkeleiden aineistosta tekoälyn käytön hyödyt diabeettisen retinopatian seulonnassa ja nämä on avattu tuloksina alaotsikoiden alle.



Kuvio 4. Tekoälyn käytön hyötyjä diabeettisen retinopatian seulontakuvauksessa.

5.1.1 Tekoälyn suoriutuminen

Useassa tutkimuksessa oli mukana tekoälyalgoritmeja, joiden **sekä herkkyys että tarkkuus** ylittivät teoreettisessa viitekehelyksessä mainitun 80 % rajan, jolloin seulonta on riittävän laadukasta (Wroblewski ym. 2023; Liljenquist ym. 2021; Shah ym. 2021; Sosale ym. 2020). Lisäksi 21 tekoälyalgoritmia vertaillaessa tutkimuksessa seitsemällä algoritmilla oli 80 % ylittävä herkkyys sekä tarkkuus, kun etsittiin jatkolähetettä vaativia kuvia (Kubin ym. 2024). Myös seitsemää tekoälyalgoritmia anonyymisti vertaillaessa tutkimuksessa yhdellä algoritmilla ylittyi sekä herkkyys, että tarkkuus 80 % (Lee ym. 2021). Muokkaamalla tekoälyalgoritmin kynnsarvoja etsimään vain jatkolähetettä vaativat kuvat muista kuvista, suoriutui tekoälyalgoritmi paremmin ja riittävällä laatutasolla (Curran ym. 2023; Meredith ym. 2023).

Yleisempää tutkimuksissa kuitenkin oli, että algoritmi saavutti riittävän **herkkyyden** (Musetti ym. 2024; Vaghefi ym. 2020) tarkkuuden kustannuksella. 80 % ylittävä herkkyys saatiin esiin seuraavilla kynnsarvojen muokkauksilla: "ei kiireinen lähete" vs. kiireinen lähete (Teoh ym. 2023), lievää suuremmat muutokset (Malerbi ym. 2022a; Wolf ym. 2021), jatkolähetettä vaativat kuvat 6 / 21 algoritmilla (Kubin ym. 2024). Myös kaksi algoritmia seitsemää algoritmia vertaillaessa tutkimuksessa ylitti herkkyydessä riittävän rajan (Lee ym. 2021). Lisäksi eräässä tutkimuksessa tekoälyalgoritmit ei jäänyt yksikään merkittävä diabeettinen retinopatia huomaamatta (Meredith ym. 2023). Tekoälyä voidaan myös hyödyntää tunnistamaan silmänpohjakuvista usein hankalasti tulkittavaa makulan alueen turvotusta. Tekoäly havaitsi riskin makulopatialle 60 silmässä, joista otettiin tarkentava OCT-kuva. Diagnoosi vahvistui 47 silmässä, jolloin tekoälyalgoritmi oli 74 % oikeassa ja tällöin vääriä positiivisia epäilyjä oli 13 kappaletta. (Malerbi ym. 2022b.)

Riittävän laadukas **tarkkuus** (mutta samalla alhaisempi herkkyys) tuli myös ilmi, kun etsittiin mitä tahansa diabeettisen retinopatian astetta (Curran ym. 2023). 80 % ylittävä tarkkuus saatiin myös esiin kynnsarvojen muokkauksilla: kun algoritmi etsi pelkästään niitä kuvia, joista ei tarvitse tehdä jatkolähetettä (Teoh ym. 2023). Kahdeksalla algoritmilla 21:sta ylittyi riittävä tarkkuus kun piti etsiä jatkolähetettä vaativia kuvia (Kubin ym. 2024). Lisäksi seitsemää tekoälyalgoritmia vertaillaessa tutkimuksessa yksi ylitti tarkkuudessa 80 % raja-arvon (Lee ym. 2021).

Tekoälyn hyviin puoliin kuuluu sen antama **nopea vastaus**. Parhaimmillaan vastaus tulee sekunneissa (Musetti ym. 2024; Vaghefi ym. 2020) tai minuutissa (Dow ym. 2023; Abramoff ym. 2023; Shah ym. 2021; Wolf ym. 2021.) Satunnaistetussa kontrolloidussa

tutkimuksessa oli 331 potilasta, jotka saivat kuvien tuloksen pelkästään tekoälyltä. Heistä 100 % oli **tyytyväisiä** vastauksen nopeuteen sekä vastauksen saamiseen tekoälyltä. (Abramoff ym. 2023.)

Tekoälyn hyötyihin kuuluu sen **kyky luokitella** parhaimmillaan jokainen sille syötetty kuva. Dataa oli toisaalta ennakkoon rajattu niin, että 183 kuvasettiä otettiin pois ihmisarvion jälkeen riittämättömän laadun vuoksi ennen tekoälyn tekemää lausuntoa. Tämä oli 1,8 % koko tutkimuksen kuvajoukkomäärästä. (Meredith ym. 2023.) 21 anonyymia tekoälyalgoritmia vertailevassa tutkimuksessa 19 algoritmia onnistui lausumaan yli 98 % kuvista (Kubin ym. 2024). Kahta tekoälyalgoritmia vertailevassa tutkimuksessa toinen algoritmeista onnistui lausumaan jokaisen kuvan, vaikkakin se arvioi 82 potilaan kohdalla kuvan huonolaatuiseksi (Wroblewski ym. 2023).

5.1.2 Tekoälyn hyviä ominaisuuksia

Hyödyllinen ominaisuus tekoälyalgoritmillä on sen toiminnan **automaattisuus**, jolloin sovellus ei vaadi ihmiseltä toimenpiteitä vastauksen saamiseksi (Musetti ym. 2024; Dow ym. 2023; Abramoff ym. 2023). Eräessä tutkimuksessa todettiin, että varsinkin maaseudulla kuvatessa, mahdollisesti huonojen verkkoyhteyksien päässä, on hyötyä siitä, että tekoälysovellus toimii ilman internetyhteyttä, **offline**, jolloin on mahdollisuus saavuttaa syrjäseudulla asuvia (Malerbi ym. 2022a). Tekoälyllä on myös mahdollisuus tasoittaa kokemattomien ja eri tasoisten kuvatulkitsijoiden eroja lausumisissa. Tekoälyn avulla lausunnot ovat tasalaatuisempia. (Teoh ym. 2023.)

Joissain tekoälyalgoritmeissa on **turvaominaisuuksia**, jotta niiden käyttö olisi turvallisempaa: Tekoälyalgoritmit havaitsevat oikean ja vasemman silmän kuvat otetuiksi, ja myös sen, onko tarkannäön alue ja papilla saatu kuvattua (Abramoff ym. 2023; Wroblewski ym. 2023; Shah ym. 2021; Vaghefi ym. 2020). Yhdessä algoritmissa oli myös ominaisuutena vastauksen ”kasvattaminen” seuraavaan luokkaan. Esimerkiksi jos kahdessa saman silmän kuvassa oli hyvin lieviä muutoksia, niin tuloksena tekoäly antoi kyseiselle silmälle vastauksen ”lieviä muutoksia” (Vaghefi ym. 2020). Yleistä oli myös ominaisuus, jossa eritasoisesti luokitellut silmät saivat potilaskohtaisen vastauksen vakavamman muutoksen mukaan (Kubin ym. 2024; Meredith ym. 2023; Vaghefi ym. 2020; Sosale ym. 2020). Myös joissain algoritmeissa tekoäly ei antanut vastausta lainkaan, jos yksikin kuva oli laadultaan riittämätön (Kubin ym. 2024; Musetti ym. 2024; Abramoff ym. 2023; Wroblewski ym. 2023; Teoh ym. 2023; Shah ym. 2021), tai siltä puuttui tarvittava määrä kuvia (Shah ym. 2021).

Hyödyllistä on myös tekoälyalgoritmin ominaisuus **ohjeistaa kuvaajaa** ottamaan uusi kuva, mikäli sen laatu on riittämätön. Tämä ominaisuus auttaa varsinkin kokemattomampia kuvaajia onnistumaan paremmin tehtävässään. (Abramoff ym. 2023; Wroblewski ym. 2023; Shah ym. 2021; Wolf ym. 2021; Vaghefi ym. 2020; Sosale ym. 2020.) Huonolaatuisesta kuvasta varoittamisen lisäksi etua on, jos sovellus varoittaa joskus hankalasti tulkittavissa olevan **makulopatian riskistä**, jolloin on mahdollisuus ottaa tarkentavia kuvia esimerkiksi OCT – kameralla (Vaghefi ym. 2020). Osa tekoälyalgoritmeista loi kuvasta ja sen havaitsemista diabeettisen retinopatian muutoksista **lämpökartan**, joka indikoi kuvaajalle ja kuvien tulkitsijalle kohdat, mitkä tekoälyn mielestä vaativat huomiota (Malerbi ym. 2022a; Malerbi ym. 2022b; Vaghefi ym. 2020).

Englantilaisessa tutkimuksessa tekoälyalgoritmi sai lausua kuvat valmistajan tehdasasetuksilla niin, että ensin käytiin läpi “ei diabeettista retinopatiaa” olevat kuvat vastaan mikä tahansa diabeettisen retinopatian aste, jonka jälkeen tekoälyalgoritmin tuli etsiä kuvista lähetettä vaativat tapaukset. Herkkyys nousi 70 prosentista 95 prosenttiin ja tarkkuus toisaalta laski 0,2 prosenttia. Tutkimuksen mukaan tekoälyalgoritmin asetusten on tärkeää olla **muokattavissa**, jotta sen **kynnysarvoja** voi vaihtaa ja täten voidaan optimoida sen toimintaa vastaamaan maassa käytössä olevaa tautiluokittelua ja taudin epidemiologisia piirteitä. Muokkaamalla kynnyksarvoja, voidaan tekoälyä myös hyödyntää erityyppisissä tehtävissä esimerkiksi apuna laaduntarkkailuun, pelkästään puhtaiden kuvien poissulkemiseen tai lähetettä vaativien kuvien löytämisessä. (Meredith ym. 2023.) Vastaavanlaisia suoritusparannuksia raportoitiin kolmessa muussakin tutkimuksessa, kun tekoälyalgoritmin kynnyksarvoja muokattiin tutkimuksen aikana etsimään eritasoisia diabeettisen retinopatian muutoksia (Curran ym. 2023; Shah ym. 2021; Sosale ym. 2020).

5.1.3 Tekoälyn käytön seurauksia

Kirjallisuuskatsauksen toisessa taloudellisessa arvioinnissa mallinnettiin 1100 erilaista herkkyys/tarkkuus paria, joista kuvajoukosta tarkimmin suoriutunut valittiin referenssiksi. Tähän referenssiin verrattuna kuusi skenaariota toi **kustannussäästöjä**. Jotta kustannussäästöjä syntyy eniten, täytyy tekoälyalgoritmin herkkyys olla yli 88 %, jolloin paritarkkuus 90 % tutkimuksen laskelmien mukaan. (Wang ym. 2024.) Seitsemää tekoälyalgoritmia vertailevassa tutkimuksessa todettiin, että kustannussäästöjä syntyy, kun käytetään parhaiten suoriutuvia algoritmeja. Arvion mukaan, säästöjä syntyisi vuosittain 1 500 000–1 800 000 dollaria, jos tekoäly lausuisi 100 000 potilaasta ei diabeettisen retinopatian tapaukset ja silmälääkäri lausuisi jäljelle jäävät huonolaatuiset kuvat, sekä diabeettisen retinopatian muutoksia sisältävät kuvat. (Lee ym. 2021.) Tekoäly myös

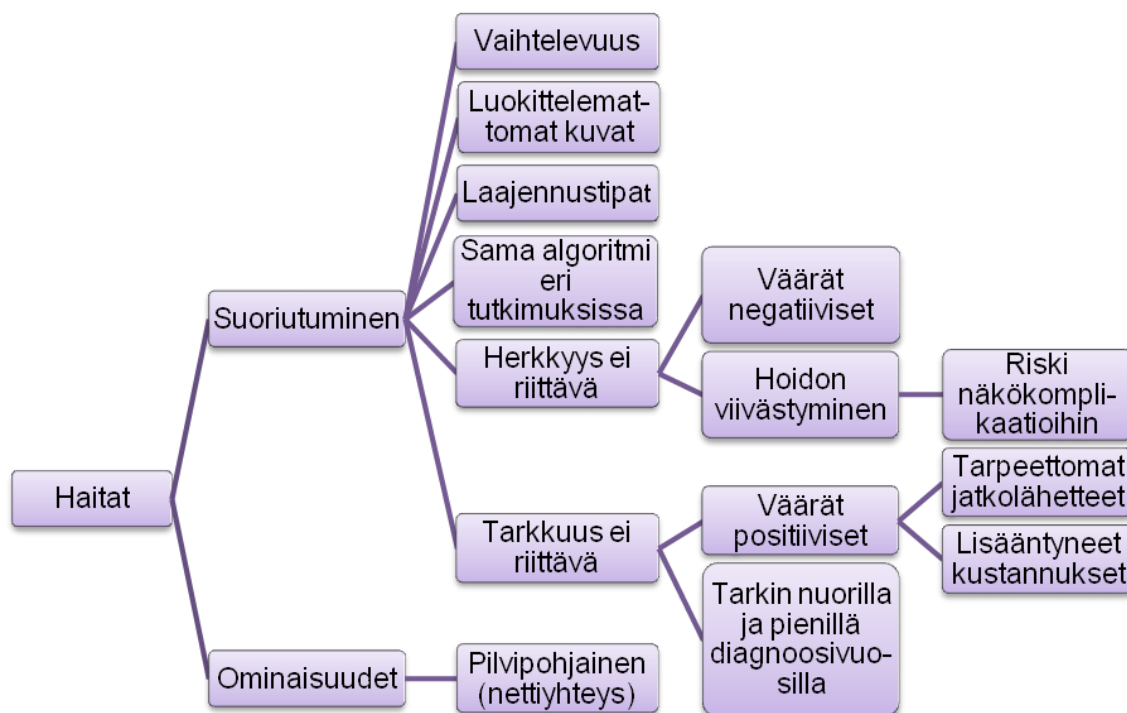
vaikuttaa kustannuksia säästävästi, koska sen käyttäminen on halvempaa, kuin silmälääkärin lausunnot (Huang ym. 2022). Kustannussäästöjen lisäksi tekoälyllä voidaan saavuttaa myös **kustannustehokkuutta**. Tutkimuksen referenssiin verrattuna seitsemän tekoälyalgoritmia loi kustannustehokkuutta, kun tarkkuus on yli 80 % ja tällöin pariterkkyys 96 % (Wang ym. 2024.) Tekoälyavusteisella seulonnalla voidaan saavuttaa väestössä enemmän laadukkaita elinvuosia verrattuna tilanteeseen, jossa silmänpohjia ei seulottaisi lainkaan tai se tehtäisiin silmälääkärin toimesta (Huang ym. 2022).

Kiinassa tehdyn taloudellisen arvioinnin mukaan tekoälystä on suuri hyöty, jos se poissulkee suuren määrän negatiivisia potilaita, ja ihmisarvioon menevät vain tapaukset, joissa on diabeettisen retinopatian muutoksia (Wang ym. 2024). Kirjallisuuskatsauksen kaikissa tutkimuksissa ei mainittu "ei diabeettista retinopatiaa" sisältävien kuvien osuutta, mutta esimerkiksi ilmoitettuja lukuja oli 59 % (Kubin ym. 2024) ja 69 % (Liljenquist ym. 2021). Otannaltaan suurimmassa tutkimuksessa arvo oli hyvin samankaltainen, kuin kahdessa edeltävässä esimerkissä. Tutkimuksessa diabeteksestä johtuvia muutoksia silmänpohjissa ei havaittu 65 %:lla kuvista. (Vaghefi ym. 2020.) Poissulkeamalla puhtaat kuvat pois, vapautuu silmälääkärin resursseja enemmän vaativiin tapauksiin, aikaa säästyy vastaanotolla sekä **tuottavuus paranee** jopa 40 %. Tutkimuksen kaikki erikoislääkärit olivat 100 % sitä mieltä, että tekoäly **säästi aikaa** heidän kliinikallansa ja myös 100 % sitä mieltä, että heiltä säästyä aikaa keskittyä enemmän niihin potilaisiin, joilla on siihen enemmän tarvetta. Tekoäly myös vähensi potilaiden odotusaikaa klinikalla, jolloin myös heidän tyytyväisyytensä saatuun palveluun paranee. (Abramoff ym. 2023.)

Kvasikokeellisessa tutkimuksessa verrattiin niiden potilaiden **osallistumista seurantaan**, jotka saivat lähetteen jatkotutkimukseen seulonnassa käymisen jälkeen. Tutkimuksessa havaittiin, että ne potilaat, jotka saivat vastauksen vain tekoälyltä, osallistuivat kolme kertaa todennäköisemmin seurantaan 90 päivän kuluessa, kuin jos vastaus olisi saatu ihmiseltä. Tekoälyltä saadun päätöksen jälkeen seurannassa kävi kolmen kuukauden kuluessa 35,5 %, kun ihmispäätöksen tai tekoälyavusteisen ihmisen tekemän päätöksen jälkeen 12 %. (Dow ym. 2023.)

5.2 Tekoälyn käytön haittoja diabeettisessa retinopatiassa

Samoin kuin tekoälyn hyödyt, myös tekoälyn haitat diabeettisen retinopatian seulontakuvauksessa koottiin kuvio 5:n tutkimusartikkeleiden taulukoinnin jälkeen. Kuvion sisältö on avattu seuraavien alaotsikoiden alle.



Kuvio 5. Tekoälyn käytön haittoja diabeettisen retinopatian seulontakuvauksessa.

5.2.1 Tekoälyn suoriutuminen

Kahdessa tutkimuksessa vertailtiin useamman anonyymin tekoälyalgoritmin suoriutumista samaan kuvajoukkoon. Toisessa tutkimuksessa oli mukana seitsemän algoritmia ja toisessa 21. Molemmissa tutkimuksissa tuli esiin algoritmien suuret erot herkkyyksien ja tarkkuuksien välillä. 21 algoritmia vertailevassa tutkimuksessa **vaihteluväli** oli 49 % - 92 % sen suhteen havaitsiko algoritmi jatkolähetettä vaativat kuvat vai ei. Tekoälyalgoritmien herkkyyksien keskiarvo oli 77,5 %, vaihdellen 13 % - 97 % väliltä. Tekoälyalgoritmien tarkkuuksien keskiarvo oli 81 %, vaihdellen 20 % - 100 % väliltä. (Kubin ym. 2024.) Seitsemän algoritmin vertailevassa tutkimuksessa herkkyydet vaihtelivat 51 % - 86 % väliltä, sekä tarkkuudet 60 % - 84 % väliltä (Lee ym. 2021). Kun verrattiin loppujen tässä kirjallisuuskatsauksessa mukana olleiden tutkimusten suoriutumista toisiinsa, oli herkkyyksien osalta vaihteluväli 70 % - 100 % (Meredith ym. 2023; Musetti ym. 2024) sekä tarkkuuksien osalta vaihteluväli 45 % - 95 % (Teoh ym. 2023; Sosale ym. 2020).

Tekoälyltä voi jäädä myös osa kuvista **luokittelematta** huonon laadun tai muun syyn vuoksi. Vaikka osassa tutkimuksia otosta oli karsittu etukäteen niin, että esimerkiksi laserhoidettuja silmänpohjia tai muuta silmänsairautta sairastavia ei otettu tutkimukseen

mukaan. Myös osassa oli kaihi poissulun syynä tai epäonnistuneet kuvat. Näissä tutkimuksissa lausumattomien kuvien osuus oli 2,5 % (Wolf ym. 2021), 3 % (Abramoff ym. 2023), 4 % (Vaghefi ym. 2020) tai jopa 13 % (Shah ym. 2021.) Tutkimuksissa, joissa otosta ei etukäteen karsittu, vaan tutkimukseen otettiin mukaan väestöä sellaisenaan kuin se kuvauksiin tuli, oli iso ero sen suhteen, miten algoritmi suoriutui: Kahta algoritmia vertailevassa tutkimuksessa lausumatta jäi 6,5 % ja 8 % kuvista (Musetti ym. 2024). Yhdysvaltalaisessa tutkimuksessa luku oli lähes 35 % (Dow ym. 2023) ja meksikolaisessa tutkimuksessa puolestaan 37 % toisella tutkimuksen algoritmeista (Wroblewski ym. 2023). Lisäksi kahdelta tekoälyalgoritmilta jäi lausumatta 11 % ja 28 % kuvista suomalaisessa tutkimuksessa (Kubin ym. 2024).

Aikuisille potilaille suunnatussa tutkimuksessa tekoälylle annettiin ensin arvioitavaksi silmänpohjakuvat, jotka oli otettu ilman **laajennustippoja**. Tekoälyalgoritmi onnistui lausumaan 87 % kuvista. Kuvauksen jälkeen niiden silmien, joista oli tullut huonolaatuiset kuvat, pupillit laajennettiin ja lausuntaprosentti kasvoi kymmenellä prosenttiyksiköllä. Tässä tutkimuksessa pupillien laajennuksella ei juurikaan ollut merkitystä tekoälyalgoritmin suoriutumiseen herkkyyden ja tarkkuuden osalta. Kun pupillit laajennettiin, kasvoi näköä uhkaavien muutosten havaitsemisen herkkyys 0,1 % ja tarkkuus 0,5 % ja lievää suurempien muutosten havaitsemisen tarkkuus 0,3 % herkkyyden pysyessä samana. (Liljenquist ym. 2021.) Puolestaan toisessa, lapsiin ja nuoriin, kohdistuneessa tutkimuksessa ei käytetty lainkaan pupilleja laajentavia silmätippoja ja tekoälyalgoritmi ylsi suoraan vastaavaan lausuntamäärään eli 97 % (Wolf ym. 2021). Lisäksi seitsemän tekoälyalgoritmin vertailututkimuksessa oli kaksi eri datajoukkoa. Toinen joukko oli kuvattu ilman laajennusta, joista huonolaatuisia oli 16 % ja puolestaan toinen joukko oli laajennetuista pupilleista ja huonolaatuisia oli 2,5 %. Tässä tutkimuksessa ei voi vetää suoraa johtopäätöstä pelkästä laajennustippojen vaikutuksesta kuvien laatuun, sillä paikassa, jossa tippoja käytettiin, oli myös parempi kuvauskoulutus ja huonoja kuvia uusintaotettiin. Laajennustippojen käyttämisellä saattoi kuitenkin olla vaikutusta huonolaatuisten kuvien suureen eroon prosentuaalisesti. (Lee ym. 2021.)

Alapuolella olevaan taulukko 6 on koottuna samaa tekoälyalgoritmi tutkineet viisi tutkimusta (Musetti ym. 2024; Dow ym. 2023; Abramoff ym. 2023; Shah ym. 2021; Wolf ym. 2021). Algoritmi on ainakin Yhdysvalloissa käytössä oleva IDx-DR, jonka nimi on nykyään LumineticsCore. Todennäköisesti sama tekoälyalgoritmi oli mukana myös kahdessa muussa tutkimuksessa, jossa vertailtiin eri algoritmien suoriutumista samaan kuvajoukkoon. Vertailevat tutkimukset olivat anonyymeja algoritmin nimien suhteen, joten niistä kahdesta ei saatu dataa tähän vertailuun (Kubin ym. 2024; Lee ym. 2021).

Taulukko 6. Tekoälyalgoritmi LumineticsCoren antamia tuloksia viidessä eri tutkimuksessa. Tulokset vaihtelevat tutkimusten välillä.

Tutkimus, vuosi, maa	Rajattu kohdejoukkoa etukäteen	Laajennustipat	Lausumattomia kuvia	Herkkyys ja tarkkuus %
Musetti ym. 2024, Italia	Ei	Ei ja kyllä	8 %	86,8 % / 60,7 % (pahin skenaario eli lausumattomat väärin) 94,3 % / 66 %
Dow ym. 2023, Yhdysvallat	Ei	Ei	34,95 %	-
Abramoff ym. 2023, Bangladesh	Kyllä	Kyllä	3,42 %	93,9 % / 84 % (arvioitu)
Shah ym. 2021, Espanja	Kyllä	Kyllä	13,2 %	100 % / 81,82 % (jatkolähte) 100 % / 94,64 % (näköä uhkaava)
Wolf ym. 2021, Yhdysvallat	Kyllä	Ei (lapsia)	2,5 %	85,7 % / 79,3 %

Taulukko 6:n viidessä tutkimuksessa oli eroja tutkimusjoukon ja etukäteen tehdyn potilasrajauksen suhteen sekä vaihtelevuutta oli myös siinä, oliko potilaiden pupilleja laajennettu tipoilla vai ei. Korrelaatiota ei näyttänyt olevan sen suhteen oliko potilasjoukkoa rajattu etukäteen tai oliko pupilleja laajennettu vs. kuinka paljon tekoälyltä jäi kuvia arvioimatta. Oletuksena oli, että mikäli pupilleja ei ole laajennettu ja tutkimusjoukkoa ei ole etukäteen rajattu, olisi lausumattomia kuvia määrällisesti enemmän, sekä tekoälyn suoriutuminen huonompaa. Myös päinvastoin, mikäli laajennusta olisi käytetty ja kohdejoukko rajattu etukäteen pitäisi oletuksena tekoälyn suoriutumisen olla parempaa. Kuten taulukosta voi havaita, **vaihtelee** tekoälyn **suoriutuminen** paljon **tutkimusten välillä**. Lausumattomia kuvia on 2,5 prosentin ja lähes 35 prosentin väliltä, eikä luku mene linjassa laajennustippojen ja kohdejoukon mahdollisen etukäteen tehdyn rajauksen kanssa. (Musetti ym. 2024; Dow ym. 2023; Abramoff ym. 2023; Shah ym. 2021; Wolf ym. 2021.) Edellä olevista seikoista voi päätellä, että tekoälyalgoritmin suoriutuminen ei ole aina tasaista ja toistuvaa, vaan siihen voi vaikuttaa esimerkiksi väestön erityispiirteet ja muun muassa kuvattavien ikäjakauma.

Alhainen algoritmin **herkkyys** kasvattaa vääriä negatiivisia tuloksia, jolloin **hoito voi viivästyä** ja tulee **riskejä näkökomplikaatioihin** (Wang ym. 2024). Algoritmin kynnysarvojen erilaisilla muokkauksilla herkkyys jäi alle 80 % tason: mikä tahansa diabeettisen retinopatian aste (Curran ym. 2023), “ei lähetettä” vaativien kuvien erottelu (Teoh ym. 2023) ja “ei diabeettista retinopatiaa” vastaan mikä tahansa diabeettisen retinopatian aste (Meredith ym. 2023). Lisäksi seitsemää tekoälyalgoritmia vertailevassa tutkimuksessa neljässä jäi herkkyys alle 80 %. Lisäksi näistä neljästä algoritmista yhdessä jäi herkkyys alimmaksi (74,5 %) tunnistamaan edenneet proliferatiiviset muutokset, joka tarkoittaa, että tekoäly ei 25,5 % havainnut edenneitä muutoksia, jolloin väärä negatiivinen tulos on riski näönmenetykseen potilaalla. (Lee ym. 2021.) Alhaisin herkkyys (13 %) löytyi suomalaisesta tutkimuksesta, kun algoritmin tuli etsiä jatkolähetettä vaativia kuvia. (Kubin ym. 2024). Tekoälyltä havaitsematta jääneet diabeettisen retinopatian muutokset olivat yleisimmin joko lieviä tai hyvin lieviä muutoksia (Liljenquist ym. 2021; Wolf ym. 2021; Vaghefi ym. 2020) tai kohtalainen diabeettinen retinopatia (Kubin ym. 2024; Malerbi ym. 2022a) tai makulan turvotus (Malerbi ym. 2022a).

Alhainen algoritmin **tarkkuus** kasvattaa vääriä positiivisia tuloksia, jolloin tulee **tarpeettomia jatkolähetetteitä ja kustannukset kasvavat** (Wang ym. 2024). Riittämättömiä, alle 80 % tarkkuuksia oli useassa tutkimuksessa eri skenaarioissa (Musetti ym. 2024; Malerbi ym. 2022a; Wolf ym. 2021; Vaghefi ym. 2020). Lisäksi eri kynnysarvojen muokkauksissa luvut saattoivat muuttua: “Ei kiireisistä” läheteteistä muokkaus tehtiin kiireisten läheteteiden etsimiseksi ja tarkkuus nousi 45 %:sta 76 % jääden silti tavoitteesta. (Teoh ym. 2023.) Lisäksi 21 tekoälyalgoritmia vertailevassa tutkimuksessa alhaisin tarkkuus yhdellä algoritmilla oli 20 % tunnistaa jatkolähetettä vaativat kuvat (Kubin ym. 2024). Tekoälyalgoritmien väärin positiivisten yleisimpiä syitä olivat muut verkkokalvosairaudet, kuten ikärappeuma ja keskuslaskimo- ja haaratukokset (Kubin ym. 2024), verkkokalvon muu patologia, joka ei liittynyt diabeettiseen retinopatiaan (Meredith ym. 2023), pigmenttiläiskät (Malerbi ym. 2022a), verkkokalvon kiiltelevä kalvo, kuvien laatuongelmat (Wolf ym. 2021) sekä artefaktit (Meredith ym. 2023; Malerbi ym. 2022a; Sosale ym. 2020). Makulaturvotuksen väärissä positiivisissa hälytyksissä yleisimmät syyt olivat kovat eksudaatit ilman turvotusta, pigmenttikertymät, kuvan artefakti (Malerbi ym. 2022b.), fovean heijaste tai drusen (Vaghefi ym. 2020).

Tekoäly oli **tarkempi luokittelemaan nuorempia potilaita**, sekä tarkkuus parani myös, kun **sairastettuja vuosia on vähemmän** (Curran ym. 2023; Shah ym. 2021). Tekoäly suoriutui paremmin keskimäärin 16-vuotiaista potilaista kuin 23-vuotiaista, sekä tarkkuus parani myös silloin, kun sairastettuja vuosia on vähemmän (5 vuotta)

kun verrattiin keskimäärin 8,5 vuotta sairastaneisiin (Curran ym. 2023). Espanjalaisessa tutkimuksessa tekoälyn tarkkuus oli 86 % jos diabeteksen diagnoosivuosisia oli kymmenen tai vähemmän. Tarkkuus laski 71 prosenttiin mikäli sairastamisvuosisia oli enemmän kuin kymmenen. Potilaiden ollessa alle 65-vuotiaita oli tarkkuus 89 % ja potilaiden iän ylittäessä 65-vuotta, tarkkuus laski 79 prosenttiin. (Shah ym. 2021.)

5.2.2 Tekoälyn ominaisuudet

Kolmessa tutkimuksessa mainittiin tekoälyalgoritmin IDx-DR:n olevan pilvipohjainen, jolloin toimiakseen täytyy olla toimiva ja nopea **nettiyhteys**. Tämä koettiin joissain määrin ongelmana, varsinkin jos on tarve kuvata seudulla, jossa infrastruktuuri ei ole kehittynyttä. (Musetti ym. 2024; Dow ym. 2023; Abramoff ym. 2023.)

6 Pohdinta

6.1 Tulosten yhteenveto

Parhaimmat tekoälyalgoritmit ovat sekä kustannustehokkaita että kustannuksia säästäviä, kun niiden herkkyys, tarkkuus ja luokittelun kynnsarvot on viillattu sopimaan mm. kunkin maan tautitilanteeseen (Wang ym. 2024; Meredith ym. 2023). Tekoälyalgoritmit pystyvät myös parhaimmillaan lausumaan automaattisesti, itsenäisesti ja nopeasti lähes jokaisen niille syötetyn kuvan (Musetti ym. 2024; Meredith ym. 2023). Potilaat ovat myös tyytyväisiä palvelun nopeuteen ja vastauksen saamiseen tekoälyltä, joka voi lisätä tyytyväisyyttä vastaanotoilla (Abramoff ym. 2023). Myös seurantakäynneillä käyvien osuus voi kasvaa (Dow ym. 2023). Tekoälyalgoritmi voi vaikuttaa tuottavuuteen säästämällä silmälääkäreiden aikaa ja resursseja vaativimpiin tapauksiin (Abramoff ym. 2023). Lisäksi algoritmi voi auttaa kokemattomampia kuvaajia onnistumaan työssään varoittamalla riittämättömästä kuvalaadusta tai varoittamalla kuvaajaa lisäkuvien tarpeesta silloin, kun on epäily makulan turvotuksesta (Vaghefi ym. 2020). Algoritmi voi myös avustaa kokemattomampia kuvalausujia tuottamaan tasalaatuisia lausuntoja (Teoh ym. 2023).

Tekoälyalgoritmien suoriutumisissa on kuitenkin paljon vaihtelua (Kubin ym. 2024; Lee ym. 2021), ja niistä heikoimpien käyttäminen ei olisi kustannustehokasta (Wang ym. 2024) tai turvallista potilaan kannalta ilman laadunvalvontaa. Myös joiltain algoritmeilta jäi suuri osa kuvista lausumatta, jolloin silmälääkäreiden lausumismäärä lisääntyisi (Mu-

setti ym. 2024; Wroblewski ym. 2023; Shah ym. 2021). Tekoälyä käytettäessä on keskeyttävä erityisesti kuvien hyvään laatuun, jolloin pupilleja laajentavien tippojen käyttämisestä on hyötyä (Liljenquist ym. 2021). Samaa algoritmia tutkimusten tuloksissa oli myös hajontaa (Musetti ym. 2024; Dow ym. 2023; Abramoff ym. 2023; Shah ym. 2021; Wolf ym. 2021), jolloin pohdittaessa tekoälyalgoritmin hankintaa suomalaisen terveydenhuoltojärjestelmään diabeettisen retinopatian seulontakuvauksessa, tulee perehtyä huolellisesti useampaan julkaistuun tutkimukseen.

6.2 Päätulosten pohdinta

Teoreettisessa viitekehysessä mainitussa kirjallisuuskatsauksessa, jossa käsiteltiin 34 tutkimusta, päädyttiin lopputulokseen, että tekoälyn suoriutuminen on hyväksyttävällä tasolla verrattuna ihmisten suoriutumiseen. Meta-analysissä luokiteltavissa olevien kuvien suhteen algoritmien keskiarvoinen herkkyys oli 94 % ja tarkkuus 89 %. (Sanil ym. 2024.) Tässä kirjallisuuskatsauksessa parhaiten suoriutuneet algoritmit, joita oli neljä kappaletta, kykenivät vastaavaan tai jopa parempaan suoritukseen kuin Sanil ym. 2024 kirjallisuuskatsauksen keskiarvoinen suoritus (Wroblewski ym. 2023; Meredith ym. 2023; Liljenquist ym. 2021; Shah ym. 2021).

Tässä kirjallisuuskatsauksessa mukana olleiden tekoälyalgoritmien suoriutumiset kuitenkin vaihtelivat laajasti, vaikka osa niistä on jo kliinisessä käytössä maailmalla. Osassa herkkyys tai tarkkuus oli niin alhainen, ettei algoritmin käyttö voi olla tarkoituksenmukaista, saati turvallista itsenäisessä käytössä (Lee ym. 2021). Variaatiota algoritmien suoriutumisen suhteen oli paljon: herkkyydet vaihtelivat 70 % - 100 % (Meredith ym. 2023; Musetti ym. 2024) välillä, sekä tarkkuudet 45 % - 95 % (Teoh ym. 2023; Sosale ym. 2020) välillä. Huonoimmat lukemat löytyivät Suomessa tehdystä tekoälyalgoritmeja anonymisti vertailevassa tutkimuksessa. Yhden algoritmin herkkyys oli vain 13 % ja toisella algoritmilla tarkkuus 20 % (Kubin ym. 2024). On hyvä kuitenkin pitää mielessä, että ihminenkin on erehtyväinen ja myös alkuperäisissä referenssiluokitteluisissa voi olla virheitä, jotka johtuivat samantyyppisistä syistä, kuin tekoälynkin virheet (Meredith ym. 2023; Vaghefi ym. 2020). Täytyy olla valmis hyväksymään, ettei mikään seulontasysteemi ole yhtä aikaa 100 % varma ja herkkä ja että virheitä sattuu kaikissa luokitteluisissa (Meredith ym. 2023). Tekoälyä voi käyttää myös avustamaan havaitsemaan silmämepohjakuvista joskus hankalasti tulkittavissa olevaa makulan turvotusta. Turvotus tarkannäönalueella voi olla vaikea erottaa ihmiselle, mutta tutkimuksissa myös tekoälyllä oli haasteita tunnistaa turvonnut makula. Väärän positiivisen hälytyksen syynä tutkimuksissa oli kovat eksudaatit ilman turvotusta, pigmenttikertymät tai kuvan artefakti (Malerbi ym. 2022b.) sekä fovean heijaste tai drusen (Vaghefi ym. 2020.)

Taulukko 6 esitellyt samaa tekoälyalgoritmia tutkineet tutkimukset myös osoittavat sen, että pelkästään yhteen tehtyyn tutkimukseen ei kannata luottaa, vaan on tehtävä mahdollisimman laaja ja tarkka taustatyö ja etukäteiselvitys olemassa olevista luotettavista tutkimuksista ennen tekoälyalgoritmin hankintapäätöstä. Rohkaisevasti muutamassa tutkimuksessa tulokset paranivat, kun algoritmin perusasetusten kynnsarvoja muokattiin löytämään jatkolähetettä vaativia tai näköä uhkaavia muutoksia (Curran ym. 2023; Meredith ym. 2023; Shah ym. 2021; Sosale ym. 2020). Kuitenkin on vaikea ajatella, että huonoiten suoriutuneista algoritmeista, joilla oli 13 % herkkyyttä tai 20 % tarkkuus (Kubin ym. 2024), saisi pelkällä säätöjen optimoinnilla toimivaa ja tarkoituksenmukaista apuvälinettä kliiniseen työhön.

Käypähoito 2024 diabeettinen retinopatia -suosituksen mukaan muutokset silmänpohjissa ovat harvinaisia ensimmäisen viiden diagnoosivuoden aikana tyypin 1 diabetesta sairastavilla. Kun sairastettuja vuosia on takana 20, on vähintään lieviä muutoksia lähes kaikilla, jotka ovat sairastuneet alta 30-vuotiaina. Näistä edenneempiä muutoksia on jo 15–50 %:lla ja makulan alueen turvotusta 10–20 %:lla. Eli tyypillistä on, että nuorilla muutoksia on vähemmän ja mitä pidempi aika diagnoosista on kulunut, sen todennäköisempää on, että diabeteksen aiheuttamia muutoksia on silmänpohjissa. (Diabeettinen retinopatia. Käypä hoito –suositus 2024.) Tutkimuksissa havaittiin, että tekoäly suoriutuu tarkemmin nuorempien potilaiden kuvista, sekä myös silloin, kun diagnoosivuosia on vähemmän (Curran ym. 2023; Shah ym. 2021). Eli tekoäly toimii tarkemmin, kun muutoksia ei ole ehtinyt tulla silmänpohjiin. Toisaalta kuitenkin tekoälyalgoritmin suoriutuminen sittenkin parani, kun tehtäväksi annettiin löytää jatkolähetettä vaativia kuvia tai näköä uhkaavia tapauksia (Curran ym. 2023; Shah ym. 2021; Sosale ym. 2020). Tekoälyn tarkkuus myös laski kynnsarvojen muokkauksesta, jolloin se pelaa ikään kuin varman päälle ja varmistaa, että kaikki potentiaaliset tautitapaukset pääsevät jatkotutkimuksiin. (Teoh ym. 2023; Sosale ym. 2020). Toisaalta eräässä tutkimuksessa algoritmin herkkyyttä tunnistaa pitkälle edenneet muutokset oli vain 74,5 % (Lee ym. 2021). Tutkimuksissa yleisimmin raportoidut syyt väriin negatiivisiin vastauksiin oli ”hyvin lieviä” tai lieviä muutoksia sisältäneitä kuvia (Wolf ym. 2021; Liljenquist ym. 2021; Vaghefi ym. 2020) tai jopa kohtalaisia muutoksia (Kubin ym. 2024; Malerbi ym. 2022a). Yhteenvetona voidaan esittää, että tekoäly on tarkimmillaan, kun potilas on nuori tai diagnoosivuosia ei ole paljon. Myös kynnsarvoja voidaan muuttaa, jolloin tekoäly toimii paremmin ja luotettavammin rajatummassa tehtävänasettelussa. Väärät negatiiviset vastaukset ovat potentiaalisesti vaarallisia potilaalle, mutta useassa tutkimuksessa tarkempi analyysi osoitti kuvissa vain vähäisempien muutosten huomaa-

matta jäämisen tekoälyn osalta. Tästä voinee päätellä, että parhaimmillaan tekoälyalgoritmi on suhteellisen turvallinen työkalu itsenäiseen kliiniseen päätöksentekoon. Laadunvarmistus on kuitenkin alkuun paikallaan potilasturvallisuuden varmistamiseksi.

Valtaosassa tutkimuksia tekoälylle annettiin arvioitavaksi kaksi tai useampi silmänpohjasta otettu kuva (Kubin ym. 2024; Musetti ym. 2024; Wang ym. 2024; Dow ym. 2023; Abramoff ym. 2023; Wroblewski ym. 2023; Meredith ym. 2023; Malerbi ym. 2022a; Liljenquist ym. 2021; Shah ym. 2021; Wolf ym. 2021; Lee ym. 2021; Vaghefi ym. 2020; Sosale ym. 2020). Suomessa diabeettisen retinopatian seulonnassa riittävä kuvausalue on määritelty kattamaan vähintään kaksi kuvaa otettuna 45 asteen kuvakentällä (Diabeettinen retinopatia. Käypä hoito –suositus 2024). Joten suurimmalla osalla tässä kirjallisuuskatsauksessa mukana olleista tekoälyalgoritmeista on sellainen kuvamäärä, jotta diabeettisen retinopatian seulonta on riittävän kattavaa. Tässä kirjallisuuskatsauksessa mukana olleiden tutkimusten suurin kuvamäärä silmää kohden oli 4–5 kuvaa meksikolaisessa tutkimuksessa, jolloin oli pyrkimys kuvata mahdollisimman suuri alue silmänpohjasta. Tutkimuksessa toisella vertailussa olleella tekoälyalgoritmilla oli ominaisuus, ettei se anna vastausta, mikäli yksikin kuva potilaan kuvista oli huonolaatuinen. Tämän algoritmin osalta lausumattomien osuus oli 37 % (Wroblewski ym. 2023.), joka oli suurin prosentuaalinen osuus lausumattomia kuvia tässä kirjallisuuskatsauksessa. Myös Uuden-Seelannin tutkimuksessa olleella algoritmilla oli turvaominaisuus, joka kasvatti silmäkohtaisen luokittelun seuraavaan tasoon, mikäli muutoksia oli nähtävissä useassa kuvassa. Turvaominaisuuden vuoksi algoritmin tarkkuus kärsi, sen ollen 61 % ja 63 % tutkimuksen kahdessa kuvasetissä. (Vaghefi ym. 2020.) Tästä voidaan vetää johtopäätös, että on arvovalinta, kuinka monta kuvaa tekoälylle annetaan lausuttavaksi, mitkä ovat tekoälyalgoritmin turvaominaisuudet huonolaatuisten kuvien ja niiden luokittelujen suhteen ja kuinka paljon siedetään turhia jatkolähetteitä ja samalla kasvaneita kustannuksia osittain näiden algoritmien turvaominaisuuksien vuoksi.

Tekoälyn käyttö seulontakuvauksessa voi kuitenkin potentiaalisesti tuoda kustannussäästöjä sekä kustannustehokkuutta. Kiinalaisessa taloudellisessa arvioinnissa päädyttiin siihen, että herkkyys olisi hyvä olla 88–96 % ja tarkkuus 80–90 %, jotta seulonta on sekä kustannustehokasta ja kustannuksia säästävää (Wang ym. 2024.) Näiden yllä olleiden rajojen väliin osui tässä kirjallisuuskatsauksessa mukana olleista tutkimuksista kahden algoritmin tulokset (Wroblewski ym. 2023; Liljenquist ym. 2021). Lukuja ei voi kuitenkaan suoraan hyödyntää kiinalaisesta tutkimuksesta Suomeen, koska kyseessä on eri maa erilaisin tautitilantein, yhteiskunnan varoin ja kuluttajahinnoin. Joitain suuntaviivoja voidaan tästä kuitenkin hyödyntää, sillä asettamalla herkkyys ja tarkkuus tar-

peeksi korkealle ja balanssiin keskenään ei tule liikaa vääriä negatiivisia, jotka ovat yksittäiselle potilaalle potentiaalisesti haitallisia, eikä toisaalta myöskään turhia vääriä positiivisia, jotka puolestaan lisäävät yhteiskunnan maksutaakkaa. Vanhassa käypähoito suosituksessa vuodelta 2014 mainittiin seulonnan olevan laadukasta, kun sen herkkyys ylittää 80 % ja tarkkuus 94 % (Diabeettinen retinopatia. Käypä hoito -suositus 2014). Kun herkkyys ja tarkkuus asetetaan riittävään balanssiin, saadaan tekoälyalgoritmista irti kaikki sen potentiaaliset hyödyt: Kustannussäästöt ja kustannustehokkuus kun herkkyys on 88–96 % ja tarkkuus 80–90 % (Wang ym. 2024.) Käyttämällä tekoälyä tulee vähemmän lausuttavia kuvia optikoille ja silmälääkäreille, jolloin resursseja säästyy vaativimpiin tapauksiin, potilaat ovat tyytyväisiä vastauksen nopeuteen (Abramoff ym. 2023) ja mahdollisesti jatkotutkimuksissa käynti tehokkaampaa (Dow ym. 2023). Yhdysvaltalaisessa kvasikokeellisessa tutkimuksessa todettiin, että tekoälyn antaman lausunnon jälkeen ihmiset kävivät kolme kertaa todennäköisemmin suositellussa seurannassa 90 päivän sisällä, kuin ihmiseltä saadun vastauksen jälkeen. Tutkimuksessa esitettiin, että syynä voi olla tekoälyn antama nopeampi vastausaika, joka oli tutkimuksessa kaksi päivää ja ihmisellä seitsemän päivää. Tässä tutkimuksessa tekoäly auttoi siinä, että potilaat kävivät todennäköisemmin seurannassa toivotussa aikaikkunassa, joka on hyvä asia sillä, mikäli seuranta ja mahdollinen hoito ei toteudu, on koko seulonta tällöin turhaa. (Dow ym. 2023.)

Kustannustehokkuuteen voi vaikuttaa myös säätämällä seulontavälit sopimaan maan tautitilanteeseen ja diabeettisen retinopatian esiintyvyyteen (Huang ym. 2022). Tämä seikka korjattiin juuri tänä vuonna Suomessa, kun kymmenen vuotta voimassa ollut Käypähoito -suositus päivitettiin muun muassa seulontaväleistään vastaamaan nykyistä taudinkuvaa (Diabeettinen retinopatia. Käypä hoito -suositus 2024). Hoitokuluihin voi vaikuttaa englantilaisen tutkimuksen arvion mukaan vähentämällä työtaakkaa tekoälyn avulla niin, että jo 50 % työkuorman vähennys auttaisi tuottamaan säästöjä. Tämä tarkoittaa sitä, että algoritmin tarkkuudeksi säädettäisiin noin 50 %, jolloin herkkyys olisi 92,4 % tasoa. (Meredith ym. 2023.) Ideaalitalanteessa tekoäly tunnistaisi tarkasti kaikki tapaukset, joissa ei ole diabeettisen retinopatian muutoksia, jolloin terveydenhuollon ammattilaisten työtaakka vähenisi noin 66 %:lla, kun lasketaan kaikkien tässä kirjallisuuskatsauksessa mukana olleiden tutkimusten tutkimusjoukosta ilmoitetut "ei diabeettista retinopatiaa" osuuksien keskiarvot: 59 % (Kubin ym. 2024) 81 % (Musetti ym. 2024) 66 % (Abramoff ym. 2023) 48 % (Wroblewski ym. 2023) 83 % (Curran ym. 2023) 15 % (Teoh ym. 2023) 73 % (Meredith ym. 2023) 63 % (Malerbi ym. 2022a) 75 % (Malerbi ym. 2022b) 69 % (Liljenquist ym. 2021) 96 % (Wolf ym. 2021) 76 % (Lee ym. 2021) 59 % (Sosale ym. 2020). Määrä olisi vieläkin suurempi, mikäli lisäksi kaikki hyvin lievät ja lievät diabeettisen retinopatian muutokset kuvissa lausuttaisiin tekoälyn avulla,

ja vain kohtalaista suuremmat muutokset ja luokittelemattomat kuvat menisivät jatkoarvioon silmälääkärille. Suomessa ensivaiheen tulkitsija saa lausua muutokset, jotka ovat korkeintaan lieviä (Diabeettinen retinopatia. Käypä hoito -suositus 2024), jolloin tekoäly vähentäisi myös ensivaiheen tulkitsijoiden työmäärää lausunnoissa, jolloin aikaa vapautuisi esimerkiksi useamman potilaan kuvaamiseen työpäivän aikana. Kuitenkin, mikäli kuvausmääriä saataisiin kasvatettua, pitäisi mahdollisesti lisääntyvän jatkohoidonkin järjestyä yhteiskunnan puolesta, tai seulontaa on turha järjestää (Malerbi ym. 2022a). Kuitenkaan, ideaalitalanne ei täysin vielä toteudu, sillä esimerkiksi Uuden-Seelannin tutkimuksen tutkimusdatasta 95–97 % arvioitiin ei näköä uhkaavaksi ja tekoälyalgoritmin arvioima luku oli 65–70 % (Vaghefi ym. 2020).

Monessa tutkimuksessa oli potilasjoukkoa rajattu etukäteen (Abramoff ym. 2023; Curran ym. 2023; Meredith ym. 2023; Malerbi ym. 2022a; Malerbi ym. 2022b; Liljenquist ym. 2021; Shah ym. 2021; Wolf ym. 2021; Lee ym. 2021; Vaghefi ym. 2020; Sosale ym. 2020). Etukäteen poisrajattujen potilaiden määrä vaihteli 1,83 % (Meredith ym. 2023) ja 24,1 % välissä (Shah ym. 2021). Syitä tehtyihin rajauksiin tutkimuksissa oli erilaisin kombinaatioin: heikkonäköisyys, aikaisempi diabeettisen retinopatian diagnoosi, aiempi annettu silmänsisäinen hoito, huono kuvalaatu, kostea rappeuma, arvet tarkannäön alueella, todettu makulaturvotus, laskimohaaratukos, silmäinjektio, laserhoito, silmäleikkaus (pl kaihi), karsastus, muu hoidossa oleva silmänsairaus, sarveiskalvosairaus ja leikkaamaton kaihi. Poissulkujen lisäksi tekoälyalgoritmilta jäi monessa tutkimuksessa kuvia lausumatta vaihtelevat määrät lukujen ollessa 0,6 % - 37 % väliä (Kubin ym. 2024; Wroblewski ym. 2023). 37 % on melko iso luku, sillä tällöin nämä kaikki tulisivat lopulta silmälääkäreiden lausuttaviksi, eikä työtaakka kevenisikään. Lisäksi näiden datarajausten vuoksi alkuperäisten tutkimusten tuloksiin täytyy suhtautua pienin varauksin, varsinkin tarkkuuden osalta.

Yhdysvaltalaisessa tutkimuksessa oli kaksi eri kuvauspaikkaa: Atlanta ja Seattle. Atlantassa käytettiin laajennustippoja ja siellä oli myös laajempi koulutus kuvaajille, jossa esimerkiksi huonoja kuvia kehoitettiin ottamaan uusiksi. Näillä metodeilla huonolaatuisia kuvia oli 2,5 %. Seattlessa kuvat otettiin ilman laajennustippoja, eikä kuvaajia koulutettu yhtä hyvin. Siellä huonolaatuisia oli 16 %. (Lee ym. 2021.) Laajennustippoja käytettäessä lausuntamäärä kasvoi toisessa yhdysvaltalaisessa tutkimuksessa 10 prosentilla (Liljenquist ym. 2021). Näiden tutkimusten perusteella voidaan suositella, että käytettäessä tekoälyä silmänpohjakuvien tulkintaan, tulisi kuvien laadun varmistamiseen käyttää aikaa, ja käyttämällä laajennustippoja saadaan suurempi osa kuvista lausuttua tekoälyn avulla.

Yhteenvedona voidaan esittää, että parhaimmillaan tekoälyalgoritmi voi olla sekä tarkka ja luotettava, sekä sen käytöstä on paljon hyödyllisiä seurauksia. Tekoälystä voi olla apua diabeettisen retinopatian seulonnan tehostamisessa, jolloin pystytään vastaamaan tulevaisuudessa kasvavaan kuvausten kysyntämäärään. Tekoälyalgoritmin hankinnassa ja valinnassa kannattaa kuitenkin olla tarkkana sekä julkisella että yksityisellä sektorilla, ettei huonot puolet estä hyötyjen täyttä potentiaalia.

6.3 Hyödynnettävyys ja jatkotutkimusaiheet

Tämän kirjallisuuskatsauksen kaikki tulokset eivät ole suoraan yleistettävissä suomalaisen terveydenhuoltojärjestelmään, mutta yleisiä suuntaviivoja voi kuitenkin hyödyntää. Esimerkiksi molemmat taloudelliset arvioinnit oli tehty Kiinassa, jossa on eroavaisuuksia Suomen tilanteeseen monelta kantilta katsoessa. Esimerkiksi Kiinassa diabeettisen retinopatian seulontakuvaus maksaa potilaalle, joka oletettavasti vaikuttaa seulonnassa käymiseen ja halukkuuteen maksaa tutkimuksesta ja täten vaikuttaa taloudelliseen arvioon (Huang ym. 2022). Toisekseen jokainen markkinoilla oleva tekoälyalgoritmi on koulutettu ja kehitetty omalla datalla. Koulutusdatan erityispiirteet ja esimerkiksi rodullinen vaihtelu vaikuttaa algoritmin suoriutumiseen eri maanosissa. Kun pohditaan tekoälysovelluksen hankintaa diabeettisen retinopatian seulontaan Suomessa, on selvitetävää laajasti eri sovellusten toimintaperiaatteet, koulutusdatat, analysoitava niistä tehtyjen tutkimuksien tuloksia ja selvitetävää algoritmien muokkauskykyä vastaamaan Suomessa käytettyä seulontaväliä. Algoritmin toimintaa olisi hyvä pystyä viilaamaan herkkyydeltään ja tarkkuudeltaan sopimaan suomalaiseen diabeettisen retinopatian epidemiologiaan erityispiirteisiin. Kustannushyöty-suhdetta pohtiessa on otettava huomioon mm. hoitojen kustannukset, lääkäreiden määrä ja bruttokansantuote. (Wang ym. 2024.) Lisäksi useampaan tutkimukseen olisi hyvä perehtyä, sillä kuten taulukosta 6 näkee, on saman tekoälyalgoritmin tutkimuksissa toisistaan poikkeavia tuloksia. Ennen algoritmin käyttöönottoa, se tulisi testata tiukoilla metodeilla todellisella käyttöympäristön aineistolla, jotta varmistetaan turvallisuus potilaskäytössä. Myös on hyvä ottaa huomioon muiden silmäsairauksien esiintyvyys Suomessa ja niiden vaikutukset diabeettisen retinopatian seulonnan onnistumiseen tekoälyn avustuksella. Yhtenä jatkotutkimusehdotuksena esitänkin sellaisten tekoälyalgoritmien suoriutumisen vertailua, joissa on mukana myös muita silmänpohjien sairauksia kuten glaukooman ja silmänpohjaurapteen havaitsemista silmänpohjakuvista.

Kvasikokeellisessa yhdysvaltalaisessa tutkimuksessa esitettiin, että potilaat kävivät tekoälyltä saamansa vastauksen jälkeen kolme kertaa todennäköisemmin suositellussa seurannassa kuin ihmiseltä saadun vastauksen jälkeen. Tutkimuksessa todettiin, että

syy saattoi olla tekoälyltä saadun vastauksen nopeudessa. Tätä asiaa ei kuitenkaan tutkimuksessa varmistettu, vain oletettiin, että syy saattoi olla vastauksen nopeudessa. (Dow ym. 2023.) Olettamisen takia pohdittiin, että tässä voisi olla mielenkiintoinen jatkotutkimusaihe: Miten potilaat suhtautuvat vastauksen saamiseen pelkästään tekoälyalgoritmilta ilman ihmisen interventiota? Luottavatko potilaat tekoälyn antamaan vastaukseen?

Yhdysvalloissa julkaistussa tutkimuksessa keskityttiin nuoriin, alle 21-vuotiaisiin DM-potilaisiin. Tekoälyalgoritmi antoi tuloksena 61 väärää positiivista lausuntoa, joiden tarkemman analyysin jälkeen 44 näytti johtuvan nuorille silmänpohjille ominaisen kalvon kiiltelystä. Algoritmin tarkkuus oli 79 %. Kuitenkaan lapsen ikä ei korreloinut väärän positiivisen todennäköisyyteen. (Wolf ym. 2021.) Puolestaan toisessa nuoriin, alle 26-vuotiaisiin, kohdistuneessa tutkimuksessa ei havaittu, että tyyppillinen kiiltely olisi vaikuttanut väärin positiivisten määriin. Tässä tutkimuksessa tarkkuus vaihteli 92–99 % väliltä sen mukaan mitä diabeettisen retinopatian vakavuusastetta algoritmin piti etsiä kuvista. (Curran ym. 2023.) Näiden kahden tutkimuksen erilaisen suoriutumisen vuoksi, ehdotetaan yhdeksi jatkotutkimusaiheeksi tutkimusta tekoälyalgoritmin suoriutumisesta lapsipotilaiden tai nuorten silmänpohjakuvista, sillä monet tekoälyalgoritmit on koulutettu aikuisten silmänpohjakuvilla tai ne on tarkoitettu käytettäväksi aikuisten silmänpohjakuvien arvioimiseen.

6.4 Eettisyys ja luotettavuus

Tähän opinnäytetyöhön ei tarvittu tutkimuslupaa, sillä kirjallisuuskatsauksessa käytetään olemassa olevaa tutkimustietoa. Kirjallisuuskatsauksessa ei myöskään tutkita ihmisiä (Kettunen 2018.), jolloin siinä ei tarvita eettistä ennakoarviointia (Eettinen ennakoarviointi 2023). Opinnäytetyösopimus kirjoitettiin Metropolia Ammattikorkeakoulun kanssa. Tutkimuseettisen neuvottelukunnan ”Hyvä tieteellinen käytäntö”-ohjeessa (HTK) annetaan ohjeet, kuinka tutkimuksissa kunnioitetaan muiden työtä ja viittaukset tehdään asianmukaisesti. Hyvän tieteellisen käytännön peruseriaatteet on avattu sisältämään myös luotettavuuden, rehellisyyden, arvostuksen ja vastuunkannon näkökulmat. (TENK 2023:12.)

Eettisyyden näkökulmasta perehdyttiin kirjallisuuskatsauksen menetelmiin asianmukaisista lähteistä. Katsauksen menetelmät ja valinnat perustettiin huolellisesti näiden tietolähteiden mukaan. Tulkinnan virhettä vähennettiin huolellisella taulukoinnilla ja muistiinpanoilla. (TENK 2023:12.) Opinnäytetyön tulokset raportoitiin rehellisesti puutteineen

päivineen (Vilkkä 2023: 99). HTK-ohjeiden mukaisesti sitouduttiin eettisyyden ja arvostuksen näkökulmasta pitämään huolta koko opinnäytetyön ajan, että lähdemerkinnät ja viittaukset tekstissä ovat oikein (TENK 2023:14.) ja myös Metropolia Ammattikorkeakoulun ohjeiden mukaiset. Opinnäytetyö tarkistettiin Turnitin plagiaatin tunnistusohjelmalla kolme kertaa työn etenemisen aikana, jolloin varmistettiin, että tuotettu teksti on itse tuotettua, eikä sitä ole kopioitu muilta tutkijoilta (TENK 2023:17).

Kirjallisuuskatsauksessa luotettavuutta ja rehellisyyttä lisää se, että koko prosessi avattiin auki aivan ensimmäisistä hakusanoista lähtien aina lopullisiin valittuihin hakusanoihin ja tietokantoihin saakka. Sisäänottokriteerit perusteltiin, löydettyistä tutkimusartikkeleista pidettiin kirjaa ja koko kirjallisuuskatsauksen haun kulku kuvattiin prisma flow -kaavioon. Näin toimien kirjallisuuskatsaus olisi mahdollisesti myöhemmin toistettavissa. (Vilkkä 2023: 12–13.) Kirjallisuuskatsauksen luotettavuutta lisää myös se, että tutkimusartikkeleiden laadut arvioitiin JBI laadunarvioiden kriteerien mukaisesti. Eettisyyttä huomioitiin laadunarvioissa niin, että ne suoritettiin arvioiden, mutta ei arvostellen (Vilkkä 2023: 99). Kirjallisuuskatsauksessa päädyttiin jättämään katsauksen ulkopuolelle kaksi tutkimusartikkelia, jotka saivat JBI-arvioinnissa alle puolet pisteistä. Toisaalta niiden mukaan ottaminen olisi voinut olla perusteltua, koska mahdollisesti ne olisivat voineet tuoda lisäarvoa tulosten yhteenvetoon. (Vilkkä 2023: 15–16.) Jättämällä nämä tutkimukset pois niiden heikkojen tutkimusasetelmien vuoksi, pyrittiin kuitenkin lopulta kasvattamaan kirjallisuuskatsauksen luotettavuutta. Eettisyyden näkökulmasta pidettiin huolta, että opinnäytetyön lopputuloksena on synteesi tutkimuksista, eikä siinä sepitetä omia tuloksia (Vilkkä 2023: 99; TENK 2023:16).

Kirjallisuuskatsauksen luotettavuutta voi heikentää se, että tiedonhaku, tutkimusartikkeleiden mukaanotto ja poissulku sekä tutkimusten laadunarviot tehtiin yhden opiskelijan toimesta (Vilkkä 2023: 74). Kuitenkin scoping -kirjallisuuskatsaus on luonteeltaan kartoittava ja kevyempi ja soveltuu näin myös yksin tehtäväksi verrattuna laajoihin aineistoihin pohjautuviin systemaattisiin katsauksiin (Vilkkä 2023: 23, 28). Tämä oli myös opinnäytetyön tekijältä ensimmäinen kirjallisuuskatsaus, joten kokemattomuus saattaa vaikuttaa. Kuitenkin apuna oli koko opinnäytetyön tekemisen ajan seminaariryhmä, jonka kokoontumisissa sai neuvoja opinnäytetyön ohjaajilta sekä opiskelijakollegoilta. Seminaariryhmän säännölliset kokoontumiset sekä opinnäytetyön ohjaajan kanssa käydyt yksilöohjaukset kasvattavat tämän kirjallisuuskatsauksen menetelmällistä onnistumista ja luotettavuutta. Lisäksi käytettiin Metropolia Ammattikorkeakoulun kirjaston informaattikkaa apuna hakulausekkeen muodostamisessa ja sopivien tietokantojen valinnassa. Informaatikon käyttö lisää kirjallisuuskatsauksen luotettavuutta ja laatua (Siltanen ym. 2023:14).

Tämän kirjallisuuskatsauksen luotettavuutta saattaa häiritä julkaisuharha, joka tarkoittaa sitä, että tehdyt alkuperäiset tutkimukset ovat eriarvoisia, kun ne valmistuvat. Tutkimukset, joiden lopputulemana on tilastollisesti merkittävä ero tutkimusryhmien välillä, julkaistaan todennäköisemmin, kuin sellaiset tutkimukset, joissa saavutettu tulos ei juuri eroa kontrolli- ja testiryhmän välillä. Tilastollisesti merkittävien tuloksien saavuttaneita tieteellisiä tutkimuksia myös tuodaan esiin enemmän, niihin viitataan enemmän ja niitä esitellään enemmän kirjoitettuina artikkeleina. Julkaisuharhan olemassaolo voi vaikuttaa kriittisesti tutkimuksiin, jotka perustuvat yksinomaan olemassa olevien tutkimuksien tutkimiseen kuten kirjallisuuskatsauksiin. (Easterbrook & Berlin & Gopalan & Matthews 1991.) On siis varsin todennäköistä, että olen tekemäni kirjallisuuskatsauksen aikana löytänyt hyviä ja päteviä tutkimusartikkeleita, mutta julkaisuharhan vuoksi niiden edustavuus on vääristynyt. Julkaisuharhaan perustuen voisi myös olettaa, että tähän kirjallisuuskatsaukseen olisi löytynyt pelkästään tutkimuksia, joissa tekoälyalgoritmit suoriutuvat täydellisesti. Kirjallisuuskatsauksessa tuli kuitenkin ilmi tekoälyalgoritmien käytöstä myös huonoja puolia, joten julkaisuharha ei ollut rajannut näitä alkuperäisiä tutkimuksia pois. On hyvä myös pitää mielessä, että löytämäni tutkimukset käsittelevät vain osan kirjallisuuskatsauksen aiheesta, eikä ole absoluuttinen totuus (Vilka 2023: 95–96). Tämän kirjallisuuskatsauksen tavoitteena oli kuitenkin kartoittaa olemassa olevaa tietoa, eikä koota täydellistä ja kattavaa synteesiä aiheesta.

Myös tutkimusten mukaanottorajauksena ollut kielirajaus voi vaikuttaa siihen, että joi-tain hyviä tutkimuksia saattoi jäädä kokonaan tarkastelematta. Löydetyt 18 tutkimusartikkeliä tässä kirjallisuuskatsauksessa tuntui kuitenkin riittävän tutkimuskysymyksiin vastaamiseen. Saavuttamattomiin artikkelin maksullisuuden vuoksi (Vilka 2023: 69,71) jäi yhteensä 23 tutkimusartikkeliä, joista suurin osa oli Taylor & Francis tietokannasta (n=21). Yllätykseksi kyseessä olevasta tietokannasta jäikin vain kaksi tutkimusartikkeliä tarkempaan analyysiin mukaan. Näin jälkikäteen pohdittuna, olisi voinut perehtyä tutkimussuunnitelman vaiheessa vielä tarkemmin saatavilla oleviin tietokantoihin, ja mennä niin sanotusti syvemmälle tutkimusartikkeleihin lupaavan otsikon ja tiivistelmän luvun jälkeen. Nyt tämän kyseisen tietokannan osalta tulos jäi laihahkoksi. Lisäksi, mikäli opinnäytetyöhön olisi ollut enemmän aikaa käytettävänä, olisi voinut käydä läpi vielä tarkemmin tutkimusartikkelien lähdeviitteet ja käyttää enemmän resursseja löytääkseen vielä enemmän tutkimusartikkeleita tähän työhön. Kuitenkin käytettävissä olleen ajan suhteen ja opinnäytetyön laajuus huomioon ottaen voi olla tyytyväinen opinnäytetyöhön päätyneiden tutkimusartikkelien lukumäärään, joka arvion mukaan on so-piva ja edustava tähän kirjallisuuskatsaukseen. Tutkimusartikkelit myös jakaantuivat maantieteellisesti eri alueille, joka on hyvä asia. Näin algoritmien suoriutuminen tuli tutkittua eri etnisillä väestöillä, joka voi vaikuttaa algoritmien suoriutumiseen. Taloudelliset

arvioinnit olivat molemmat tehty Kiinassa. Tästä aiheesta olisi ollut mielenkiintoista käsitellä myös länsimaissa tehty tutkimus. Kiinassa tehdyt taloudelliset tutkimukset voivat kuitenkin antaa suuntaa tekoälyalgoritmien taloudellisuudesta myös muualla maailmassa.

Monessa tutkimusartikkelissa oli rajattu pois potilasjoukosta mahdollisesti huonolaatuisia kuvia. Tämä voi johtua esimerkiksi siitä, että algoritmin valmistaja on rakentanut algoritmin toimimaan tietyllä tavalla. Tämä rajaus saattaa vaikuttaa tutkimusten luotettavuuteen antamalla parempia tuloksia. Jos opinnäytetyö tehtäisiin uudelleen, voisi yrittää ottaa tutkimusten mukaanottokriteereissä huomioon sen, että mukaan tulisi pelkääntään tutkimuksia, joissa dataa ei ole rajattu etukäteen. Mahdollisesti tuon rajauksen lisäämisen vastapainoksi olisi pitänyt ottaa mukaan enemmän tietokantoja, jotta tutkimusartikkeleita löytyy varmasti riittävä määrä.

Opinnäytetyön viimeistelyvaiheessa käytiin tarkastelemassa tietokantoja uudelleen kirjallisuuskatsauksen hakulausekkeella ja tietokantojen hakurajauksilla (taulukko 2). Jostain syystä Medic ja Cinahl Complete tietokannat antoivat erilaiset hakutulokset lukumääriltään. Esimerkiksi Cinahlista löytyi 162 artikkelia vähemmän ennen tarkempia hakurajauksia, ja 55 artikkelia vähemmän hakurajauksen jälkeen. Syytä ei löytynyt, miksi näin on tapahtunut. Mahdollisesti syynä saattaa olla tietokantojen päivitykset siellä julkaistavien artikkelien suhteen. Tämä on harmillinen seikka, sillä kirjallisuuskatsaus ei ole välttämättä täysin toistettavissa tietokantojen erilaisen toimivuuden vuoksi. Taylor & Francis Onlinen ja Sage Journalsin suhteen hakulauseke kuitenkin toi rajauksen jälkeen samat tulokset, mutta 1. kesäkuuta tehdyn tietokantahaun jälkeen oli jo julkaistu joitain uusia tutkimusartikkeleita, joista pelkän otsikon perusteella olisi saattanut sopia tähän kirjallisuuskatsaukseen seitsemän tutkimusartikkelia. Uusia tutkimuksia siis julkaistaan jatkuvasti, joten tieto tekoälystä ja sen mahdollisesta potentiaalista diabeettisen retinopatian seulontakuvauksessa kasvaa jatkuvasti.

Lähteet

Abramoff, Michael D. ym. 2023. Autonomous artificial intelligence increases real-world specialist clinic productivity in a cluster-randomized trial. *NPJ Digital Medicine* 6 (1). 184.

Curran, Katie ym. 2023. CHILdSTAR: CHILdren Living With Diabetes See and Thrive with AI Review. *Clinical Medicine Insights: Endocrinology and Diabetes* 16.

Eettinen ennakkoarviointi 2023. Tutkimuseettinen neuvottelukunta (TENK). Päivitetty 10.10.2023. <<https://tenk.fi/fi/eettinen-ennakkoarviointi>> Viitattu 30.1.2024.

Easterbrook, Philippa J. & Berlin, Jesse A. & Gopalan, Ramana & Matthews, David R. 1991. Publication bias in clinical research. *The Lancet* 337 (8746). 867–872.

Diabeettinen retinopatia. Käypä hoito -suositus 2014. Suomalaisen Lääkäriseuran Duodecimin, Suomen Silmälääkäriyhdistyksen ja Diabetesliiton lääkarineuvoston asettama työryhmä. Helsinki: Suomalainen Lääkäriseura Duodecim. <www.käypähoito.fi> Viitattu 6.2.2024.

Diabeettinen retinopatia. Käypä hoito -suositus 2024. Suomalaisen Lääkäriseuran Duodecimin, Suomen Silmälääkäriyhdistyksen ja Diabetesliiton lääkarineuvoston asettama työryhmä. Helsinki: Suomalainen Lääkäriseura Duodecim. <www.käypähoito.fi> Viitattu 18.9.2024.

Diabetes. Käypä hoito -suositus 2018. Suomalaisen Lääkäriseuran Duodecimin, Suomen sisätautilääkäreiden yhdistyksen ja Diabetesliiton lääkarineuvoston asettama työryhmä. Helsinki: Suomalainen Lääkäriseura Duodecim. <www.käypähoito.fi> Viitattu 5.2.2024.

Diabetesliitto 2021. Yleistä diabeteksestä. Päivitetty 9.11.2021. <https://www.diabetes.fi/diabetes/yleista_diabeteksesta#a495674c> Viitattu 10.3.2024.

Dow, Eliot R. ym. 2023. Artificial Intelligence Improves Patient Follow-Up in a Diabetic Retinopathy Screening Program. *Clinical Ophthalmology* 17. 3323–3330.

Elements of AI. Verkkokurssi. Helsingin yliopisto ja MinnaLearn. <<https://course.elementsofai.com/fi/1>> Viitattu 10.3.2024.

FDA = U.S. Food and Drug Administration.

FDA 2018. FDA permits marketing of artificial intelligence-based device to detect certain diabetes-related eye problems. U.S. Food & Drug Administration. <<https://www.fda.gov/news-events/press-announcements/fda-permits-marketing-artificial-intelligence-based-device-detect-certain-diabetes-related-eye>> Viitattu 2.2.2024.

Gawęcki, Maciej 2021. Diabetic Retinopathy. Ranska: Laboratoire Théa.

Gospic, Katarina A.M. & Passmore, Greg 2022. Artificial Intelligence in Medicine: Importance of Ai in Medicine. E-kirja.

Hakama, Matti & Malila, Nea 2008. Millainen on hyvä seulonta? Lääketieteellinen Aikakauskirja Duodecim 124 (19). 2193–2199. <<https://www.duodecimlehti.fi/duo97538>> Viitattu 8.2.2024.

Hautala, Nina 2020. Silmätautialan sote-tavoitteet. Suomen silmälääkäriyhdistys. <<https://siljalaakariyhdistys.fi/kannanotot/silmatautialan-sote-tavoitteet/>> Viitattu 13.12.2024.

Huang, Xiao-Mei ym. 2022. Cost-effectiveness of artificial intelligence screening for diabetic retinopathy in rural China. BMC health services research 22 (1). 260.

Huhtanen, Heidi & Nyman, Mikko & Karlsson, Antti & Hirvonen, Jussi 2020. Tekoäly radiologiassa. Lääketieteellinen Aikakauskirja Duodecim 136 (17). 1957–1964. <<https://www.duodecimlehti.fi/duo15753>> Viitattu 23.3.2024.

HUS 2024. Hoitoonpääsy erikoissairaanhoidossa. <<https://www.hus.fi/potilaalle/opas-potilaalle/oikeutesi-potilaana/hoitoonpaasy-erikoissairaanhoidossa>> Viitattu 13.12.2024.

IDF = International Diabetes Federation.

IDF 2015. International Diabetes Federation and the Fred Hollows Foundation. Diabetes eye health: A guide for health care professionals <<https://idf.org/about-diabetes/diabetes-complications/eyes/>> Viitattu 8.2.2024.

IDF 2019. International Diabetes Federation. Clinical Practise Recommendations for Managing Diabetic Macula Edema. <<https://idf.org/media/uploads/2019/09/IDF-DME-CPR.pdf>> Viitattu 8.2.2024.

IDF 2024. International Diabetes Federation. <<https://idf.org/>> Viitattu 8.2.2024.

Ilanne-Parikka, Pirjo 2021. Diabetes (“sokeritauti”). Duodecim Terveyskirjasto 124 (19). 2193–2199. <<https://www.terveyskirjasto.fi/dlk00011>> Viitattu 5.2.2024.

Jalovaara, Marika ym. 2023. Väestörakenteen muutos haastaa yhteiskunnan kestävyysden. Väestöliitto. <https://storage.googleapis.com/vaestoliitto-production/2023/01/afd88237-demography_tietopaketti_digi_sivuittain.pdf> Viitattu 10.3.2024.

Kettunen, Jyrki 2018. Artikkel. Selvitä, tarvitsetko tutkimuksellesi luvan. Vastuullinen tiede. Päivitetty 24.7.2019. <<https://vastuullinentiede.fi/fi/tutkimuksen-suunnittelu/selvitä-tarvitsetko-tutkimuksellesi-luvan>> Viitattu 30.1.2024.

Kirkonpelto, Tia-Maria & Mäntyranta, Taina 2023. Sosiaali- ja terveysministeriön julkaisu 2023:8. Tiekartta 2022–2027. Sosiaali- ja terveysalan henkilöstön riittävyyden ja

saatavuuden turvaaminen. <<https://julkaisut.valtioneuvosto.fi/handle/10024/164634>> Viitattu 4.4.2024.

Kolari, Jukka & Kallio, Aleksi 2023. Tekoäly 123 Matkaopas tulevaisuuteen. Jyväskylä: Docendo.

Kubin, Anna-Maria & Huhtinen, Petri & Ohtonen, Pasi & Keskitalo, Antti & Wirkkala, Joonas & Hautala, Nina 2024. Comparison of 21 artificial intelligence algorithms in automated diabetic retinopathy screening using handheld fundus camera. *Annals of medicine* 56 (1).

Lee, Aaron Y. ym. 2021. Multicenter, Head-to-Head, Real-World Validation Study of Seven Automated Artificial Intelligence Diabetic Retinopathy Screening Systems. *Diabetes Care* 44 (5). 1168–1175.

Lehtonen, Olli-Pekka & Saari, Samuli & Kinnunen, Juha & Kinnunen, Marina 2023. Johtaminen hyvinvointialueella. Miten ratkaista henkilöstöpula? E-kirja. Helsinki: Alma Talent.

Liljenquist, David ym. 2021. Pivotal Evaluation of an Artificial Intelligence System for Autonomous Detection of Referrable and Vision-Threatening Diabetic Retinopathy. *JAMA network open* 4 (11).

Malerbi, Fernando Korn ym. 2022a. Diabetic Retinopathy Screening Using Artificial Intelligence and Handheld Smartphone-Based Retinal Camera. *Journal of diabetes science and technology* 16 (3). 716–723.

Malerbi, Fernando Korn & Mendes, Giovana & Barboza, Nathan & Morales, Paulo Henrique & Montargil, Roseanne & Andrade, Rafael Ernane 2022b. Diabetic Macular Edema Screened by Handheld Smartphone-based Retinal Camera and Artificial Intelligence. *Journal of medical systems* 46 (1). 8.

Meredith, Sarah & van Grinsven, Mark & Engelberts, Jonne & Clarke, Dominic & Prior, Vicki & Vodrey, Jo & Hammond, Alison & Muhammed, Raja & Kirby, Philip 2023. Performance of an artificial intelligence automated system for diabetic eye screening in a large English population. *Diabetic medicine: A journal of the British Diabetic Association* 40 (6).

Metropolia Libguides 2024a. Tiedonhaun ABC. Hakutekniikka. <<https://libguides.metropolia.fi/tiedonhaku/tiedon-hakeminen>> Viitattu 16.2.2024.

Metropolia Libguides 2024b. Tiedonhaku sosiaali- ja terveysalalla. Tietokannat ja e-aineistot. <<https://libguides.metropolia.fi/sotealat/tietokannat>> Viitattu 16.2.2024.

Mokhashi, Nikita & Grachevskaya, Julia & Cheng, Lorrie & Yu, Daohai & Lu, Xiaoning & Zhang, Yi & Henderer, Jeffrey D. 2022. A Comparison of Artificial Intelligence and Human Diabetic Retinal Image Interpretation in an Urban Health System. *Journal of Diabetes Science and Technology*. 16 (4). 1003–1007.

Musetti, Donatella & Cutolo, Carlo Alberto & Bonetto, Monica & Giacomini, Mauro & Maggi, Davide & Viviani, Giorgio Luciano & Gandin, Ilaria & Traverso, Carlo Enrico & Nicolò, Massimo 2024. Autonomous artificial intelligence versus teleophthalmology for diabetic retinopathy. *European journal of ophthalmology*.

Mäkelä, Marjukka & Varonen, Helena & Teperi Juha 1996. Systemoitu kirjallisuuskatsaus tiedon tiivistäjänä. *Lääketieteellinen Aikakauskirja Duodecim* 112 (21). 1999–
<<https://www.duodecimlehti.fi/duo60413>> Viitattu 1.2.2024.

Pajunpää, Hannu 1999. Diabeettisen retinopatian valokuvaseulonnan kustannukset ja hyödyt sekä näkövammaisten elämänlaatu ja kuolleisuus. Oulun yliopisto. OuluREPO – Oulun yliopiston julkaisuarkisto. <<https://urn.fi/URN:ISBN:9514252470>> Viitattu 13.10.2024.

Pitkämä, Marianne 2024. Kirjallisuuskatsaukset YAMK opinnäytetöissä. Luento 25.1.2024. Viitattu 3.3.2024.

PRISMA 2024. Prisma transparent reporting of systematic reviews and meta-analyses. PRISMA 2020 flow diagram for new systematic reviews which included searches of databases and registers only. <<http://prisma-statement.org/prismastatement/flowdiagram.aspx?AspxAutoDetectCookieSupport=1>> Viitattu 15.2.2024.

Sahlsten, Jaakko & Jaskari, Joel & Kaski, Kimmo & Hietala, Kustaa 2020. Diabeettisen retinopatian ja makulaturvotuksen luokittelu syväoppivan tekoälyjärjestelmän avulla. *Lääketieteellinen Aikakauskirja Duodecim* 136 (17). 1971–1978.

Sanil, Joseph & Jerrome, Selvaraj & Iswarya, Mani & Thandavarayan, Kumaragurupari & Xianwen, Shang & Poonam, Mudgil & Thulasiraj, Ravilla & Mingguangm He 2024. Diagnostic accuracy of artificial intelligence based automated diabetic retinopathy screening in real-world settings: a systematic review and meta-analysis. *American Journal of Ophthalmology*. *American Journal of Ophthalmology* 263. 214–230.

Schleutker, Elina 2013. Väestön ikääntyminen ja hyvinvointivaltio. Mitä vaihtoehtoja meillä on? *Julkari. Yhteiskuntapolitiikka* 78 (4). <<https://www.julkari.fi/bitstream/handle/10024/110779/schleutker.pdf?sequence>> Viitattu 23.3.2024.

Seppänen, Matti 2021. Diabeteksen silmäsairaus (diabeettinen retinopatia). *Duodecim Terveyskirjasto*. <<https://www.terveyskirjasto.fi/dlk00826/diabeteksen-silmasairaus-diabeettinen-retinopatia>> Viitattu 5.2.2024.

Shah, Abhay ym. 2021. Validation of Automated Screening for Referable Diabetic Retinopathy With an Autonomous Diagnostic Artificial Intelligence System in a Spanish Population. *Journal of diabetes science and technology* 15 (3). 655–663.

Siltanen, Hannele & Hamari, Lotta & Heikkilä, Kristiina & Marin, Kaisa & Parisod, Heidi & Holopainen, Arja 2023. Hoitosuosituksen laadinta - käsikirja suositustyöryhmille. Versio 3.0. Helsinki: Hoitotyön tutkimussäätiö. <<https://hotus.fi/wp-content/uploads/2023/10/hoitosuosituskasikirja-30.pdf>> Viitattu 5.3.2024.

Suistola, Anna 2024. OYS tekee teköälystä arkea hoitotyössä - Diabetespotilas saa diagnoosin minuuteissa. Kaleva. Artikkel. <<https://www.kaleva.fi/oys-tekee-tekoalysta-arkea-hoitotyossa-diabetespot/6177423#kommentit>> Viitattu 25.2.2024.

Sosale, Bhavana & Sosale, Aravind R & Murthy, Hemanth & Sengupta, Sabyasachi & Naveenam, Muralidhar 2020. Medios- An offline, smartphone-based artificial intelligence algorithm for the diagnosis of diabetic retinopathy. Indian Journal of Ophthalmology 68 (2). 391–395.

Summanen, Paula & von Wendt, Gunvor 2014. Diabetesta sairastavien potilaiden silmänpohjaseulonnan laatumittari. Käypä hoito. <<https://www.kaypahoito.fi/nix00505>> Viitattu 10.2.2024.

Suomen Lääkäriliitto 2021. Seulontatutkimukset terveydenhuollossa. <<https://www.laakariliitto.fi/laakarin-etiikka/terveyden-edistaminen-ja-sairauksien-ennaltaehkaisy/seulontatutkimukset-terveydenhuollossa/>> Viitattu 26.9.2024.

Tarnanen, Kirsi & Summanen, Paula & Komulainen, Jorma 2024. Diabeettinen retinopatia. Käypä hoito potilasversio. Julkaistu 9.8.2024 <<https://www.kaypahoito.fi/khp00059>> Viitattu 2.10.2024.

TENK = Tutkimuseettinen neuvottelukunta.

TENK 2023. Tutkimuseettinen neuvottelukunta. HTK-ohje: Hyvä tieteellinen käytäntö ja sen loukkausepäilyjen käsitteleminen Suomessa 2023. <https://tenk.fi/sites/default/files/2023-03/HTK-ohje_2023.pdf> Viitattu 30.1.2024.

Teoh, Chin Sheng ym. 2023. Variability in Grading Diabetic Retinopathy Using Retinal Photography and Its Comparison with an Automated Deep Learning Diabetic Retinopathy Screening Software. Healthcare 11.

Terveyskylä 2023. Diabeettinen retinopatia ja raskaus. Päivitetty 11.4.2023 <<https://www.terveyskyla.fi/silmasairaudet/tietoa/raskaus-ja-silmat/diabeettinen-retinopatia-ja-raskaus>> Viitattu 18.9.2024.

THL = Terveyden- ja hyvinvoinnin laitos.

THL 2023. Terveyden- ja hyvinvoinnin laitos. Diabetes. Päivitetty 7.12.2023. <<https://thl.fi/aiheet/kansantaudit/diabetes>> Viitattu 5.2.2024.

Toivonen, Hannu 2023. Mitä tekoäly on? 100 kysymystä ja vastausta. E-kirja. Helsinki: Kustannusosakeyhtiö Teos.

Tolkkinen, Laura 2022. Näkövammarekisterin vuosikirja 2022. Terveyden ja hyvinvoinnin laitos ja Näkövammaisten liitto ry. <<https://www.nakovammaistenliitto.fi/fi/nakovammarekisteri>> Viitattu 10.2.2024.

Tuomi, Jouni & Sarajärvi, Anneli 2018. Laadullinen tutkimus ja sisällönanalyysi. E-kirja. Kustannusosakeyhtiö Tammi.

Vaghefi, E. & Yang, S. & Xie, L. & Hill, S. & Schmiedel, O. & Murphy, R. & Schmiedel, D. 2020. THEIA™ development, and testing of artificial intelligence-based primary triage of diabetic retinopathy screening images in New Zealand. *Diabetic medicine: A journal of the British Diabetic Association* 38 (4).

Vilkkä, Hanna 2023. Kirjallisuuskatsaus metodina, opinnäytetyön osana ja tekstilajina. Tallinna: Printon.

Von Wendt, Gunvor 2014. Silmänpohjavalokuvauksen edut. Käypä hoito. <<https://www.kaypahoito.fi/nix00506>> Viitattu 10.2.2024.

Von Wendt, Gunvor & Komulainen, Jorma 2014. Retinopatiaseulonnan kustannushyötysuhde. Käypä hoito. <<https://www.kaypahoito.fi/nak05297>> Viitattu 10.3.2024.

Wang, Yueye ym. 2024. Economic evaluation for medical artificial intelligence: accuracy vs. cost-effectiveness in a diabetic retinopathy screening case. *Npj Digital medicine* 7 (43).

WHO = World Health Organization.

WHO 2024. World Health Organization. Diabetes. <https://www.who.int/health-topics/diabetes?gclid=EAlaIQobChMlvf-H14ybhAMVOVNBAh2-pwC-NEAAYASAAEgLrrfD_BwE#tab=tab_1> Viitattu 8.2.2024.

Wolf, Risa M. & Liu, Alvin & Thomas, Chrystal & Prichett, Laura & Zimmer-Galler, Ingrid & Smith, Kerry & Abramoff, Michael D. & Channa, Roomasa 2021. The SEE Study: Safety, Efficacy, and Equity of Implementing Autonomous Artificial Intelligence for Diagnosing Diabetic Retinopathy in Youth. *Diabetes Care* 44 (3). 781–787.

Wolf, Risa M. & Channa, Roomasa & Abramoff, Michael D. & Lehmann, Harold P. 2020. Cost-effectiveness of Autonomous Point-of-Care Diabetic Retinopathy Screening for Pediatric Patients With Diabetes. *Jama Ophthalmology*. 138 (10). 1063–1069.

Wroblewski, John J & Sanchez-Buenfi, Ermilo I & Inciarte, Miguel & Berdia, Jay & Blake, Lewis & Wroblewski, Simon & Patti, Alexandria & Suter, Gretchen & Sanborn, George E 2023. Diabetic Retinopathy Screening Using Smartphone-Based Fundus Photography and Deep-Learning Artificial Intelligence in the Yucatan Peninsula: A Field Study. *Journal of diabetes science and technology*.