

TARTU ÜLIKOOL
LOODUS- JA TÄPPISTEADUSTE VALDKOND
MOLEKULAAR- JA RAKUBIOLOOGIA INSTITUUT
GENOOMIKA INSTITUUT

Valentina Rukins

**Preeklampsia ja rasedusaegse hüpertensiooni geneetika ning seos
teiste diagnoosidega TÕ Eesti Geenivaramu ja FinnGen andmetel**

Magistritöö

30 EAP

Juhendajad kaasprofessor Triin Laisk
professor Reedik Mägi

TARTU 2023

Infoleht

Preeklampsia ja rasedusaegse hüpertensiooni geneetika ning seos teiste diagnoosidega TÜ Eesti Geenivaramu ja FinnGen andmetel

Kokkuvõte. Hüpertensiivsed raseduskomplikatsioonid, sealhulgas preeklampsia ja rasedushüpertensioon, on sageduselt teine rasedate surmapõhjus maailmas. Lisaks tõsistele lühiajalistele tüsistustele, võivad need raseduskomplikatsioonid negatiivselt mõjutada nii ema kui ka last hilisemas elus. Vaatamata selles vallas tehtavatele uuringutele, on preeklampsia täpsed bioloogilised mehhanismid jäänud ebaselgeks. Käesolvas magistritöös annoteeriti uuritavate haigustega seotud geneetilisi variante, mis tuvastati Eesti Geenivaramu ning FinnGen andmete meta-analüüsil. Lisaks hinnati SNP-põhist päritavust ning geneetilist korrelatsiooni preeklampsia, rasedushüpertensiooni ja mitme teise diagnoosi/tunnuse vahel. Töö tulemused toetavad teooriat, mille kohaselt preeklampsia ja rasedushüpertensioon jagavad geneetilist tausta.

Märksõnad: preeklampsia, rasedushüpertensioon, GWAS, LD skoori regressioon

CERCS kood: B110 Bioinformaatika, meditsiiniinformaatika, biomatemaatika, biomeetrika; B790 Kliiniline geneetika

The genetics of preeclampsia and gestational hypertension, and association with other diagnoses in the Estonian Biobank and FinnGen data

Abstract. Hypertensive disorders of pregnancy, including preeclampsia and gestational hypertension, are the second leading cause of maternal deaths worldwide. In addition to severe short-term complications, they can negatively affect both mother and the child long-term. Despite the efforts made, biological mechanisms underlying preeclampsia remain poorly understood. In this master's thesis, we annotated the genetic variants associated with the investigated disorders in a joint meta-analysis of Estonian Biobank and FinnGen data. Furthermore, we estimated the SNP-based heritability and genetic correlation between preeclampsia, gestational hypertension, and numerous other diagnoses/traits. The results of this work support the idea that preeclampsia and gestational hypertension have a shared genetic background.

Keywords: preeclampsia, gestational hypertension, GWAS, LD score regression

CERCS code: B110 Bioinformatics, medical informatics, biomathematics, biometrics; B790 Clinical genetics

Sisukord

Infoleht.....	2
Sisukord.....	3
Kasutatud lühendid.....	5
Sissejuhatus	6
1. Kirjanduse ülevaade	7
1.1. Preeklampsia.....	7
1.2. Rasedusaegne hüpertensioon.....	10
1.3. Ülegenoomsed assotsiatsiooniuuringud	11
1.4. Preeklampsia ja rasedushüpertensiooni senised ülegenoomsed seoseuuringud.....	14
2. Eksperimentaalosa	18
2.1. Töö eesmärgid.....	18
2.2. Materjalid ja meetodika	18
2.2.1. Valim ja uuritavad tunnused	18
2.2.2. Genotüpiseerimine, kvaliteedikontroll, imputatsioon.....	19
2.2.3. Assotsiatsioonianalüüs TÜ Eesti Geenivaramu ja FinnGen R7 andmetel.....	20
2.2.4. GWAS metaanalüüs TÜ Eesti Geenivaramu ja FinnGen 7R andmetel.....	21
2.2.5. GWAS metaanalüüsi tulemuste annoteerimine ja visualiseerimine FUMA veebiplatvormi abil ..	21
2.2.6. LD skoori regressioon.....	23
2.3. Tulemused	26
2.3.1. Geneetiliste variantide prioritseerimine FUMA-s	26
2.3.2. Kandidaatgeenide prioritseerimine FUMA-s.....	30
2.3.3. LD skoori regressioon.....	35
2.4. Arutelu.....	41
Kokkuvõte	44
Summary.....	45
Tänuõnad.....	47
Kasutatud kirjandus	48
Kasutatud veebiaadressid	54
Lisad	55

Lisa 1.....	55
Lisa 2.....	56
Lisa 3.....	62
Lisa 4.....	63
Lisa 5.....	64

Kasutatud lühendid

CI	usaldusintervall (ingl <i>confidence interval</i>)
CTG–VL	komplekstunnuste geneetika virtuaallabor (ingl <i>Complex Traits Genetics Virtual Lab</i>)
EAF	efektialleeli sagedus (ingl <i>effect allele frequency</i>)
eQTL	ekspressiooni taset mõjutav lookus (ingl <i>expression quantitative trait loci</i>)
FUMA	funktsionaalne kaardistamine ja annoteerimine (ingl <i>Functional Mapping and Annotation</i>)
GH	rasedushüpertensioon (ingl <i>gestational hypertension</i>)
GWAMA	ülegenoomsete seoseuringute metaanalüüs (ingl <i>Genome-Wide Association Meta Analysis</i>)
GWAS	ülegenoomne seoseuring (ingl <i>genome-wide association study</i>)
h^2	päritavuskoefitsient (ingl <i>heritability</i>)
kb	tuhat aluspaari (ingl <i>kilobase pair</i>)
LD	ahelduse tasakaalutus (ingl <i>linkage disequilibrium</i>)
MAF	harvema alleeli sagedus (ingl <i>minor allele frequency</i>)
MAGMA	geenipõhise analüüsi meetod (ingl <i>Multi-marker Analysis of GenoMic Annotation</i>)
O13	RHK-10 tähtnumbriline kood, rasedushüpertensioon
O14	RHK-10 tähtnumbriline kood, preeklampsia
PE	preeklampsia (ingl <i>preeclampsia</i>)
r_g	geneetiline korrelatsioon (ingl <i>genetic correlation</i>)
RHK-10	Rahvusvaheline Haiguste Klassifikatsioon, 10. väljaanne (ingl <i>International Classification of Diseases 10th Revision</i>)
SD	standardhälve (ingl <i>standard deviation</i>)
SE	standardviga (ingl <i>standard error</i>)
sEng	lahustuv endogliin (ingl <i>soluble endoglin</i>)
sFlt-1	FMS-sarnase türosiini kinaas-1 lahustuv retseptor (ingl <i>soluble fms-like tyrosine kinase-1</i>)
SNP	üksiknukleotiidne polümorfism (ingl <i>single nucleotide polymorphism</i>)
TÜ EGV	Tartu Ülikooli Eesti Geenivaramu

Sissejuhatus

Preeklampsia ja rasedushüpertensioon kuuluvad hüpertensiivsete raseduskomplikatsioonide hulka, mis on sageduselt teine surmapõhjus rasedate seas, moodustades 14% kõigist rasedate surmadest (Say *et al.*, 2014).

Preeklampsia ja rasedushüpertensioon jagavad mitmeid riskifaktoreid nagu näiteks esimene rasedus, mitmikrasedus, suhkurtõbi ja kõrgeenenud kehamassiindeks. Pealegi on viimases raseduses preeklampsia esinemine rasedushüpertensiooni riskifaktoriks ja vastupidi (Shen *et al.*, 2017; ülevaade Dimitriadis *et al.*, 2023). Samuti on varasematest uuringutest selgunud, et ligikaudu neljandikul rasedushüpertensiooniga naistest kujuneb välja preeklampsia (Saudan *et al.*, 1998). Kuigi preeklampsia ja rasedushüpertensiooni sümptomid on kohati sarnased (süstoolne vererõhk ≥ 140 mmHg ja/või diastoolne vererõhk ≥ 90 mmHg), kas need on eraldiseisvad tüsistused või sama spektri erinevad väljundid.

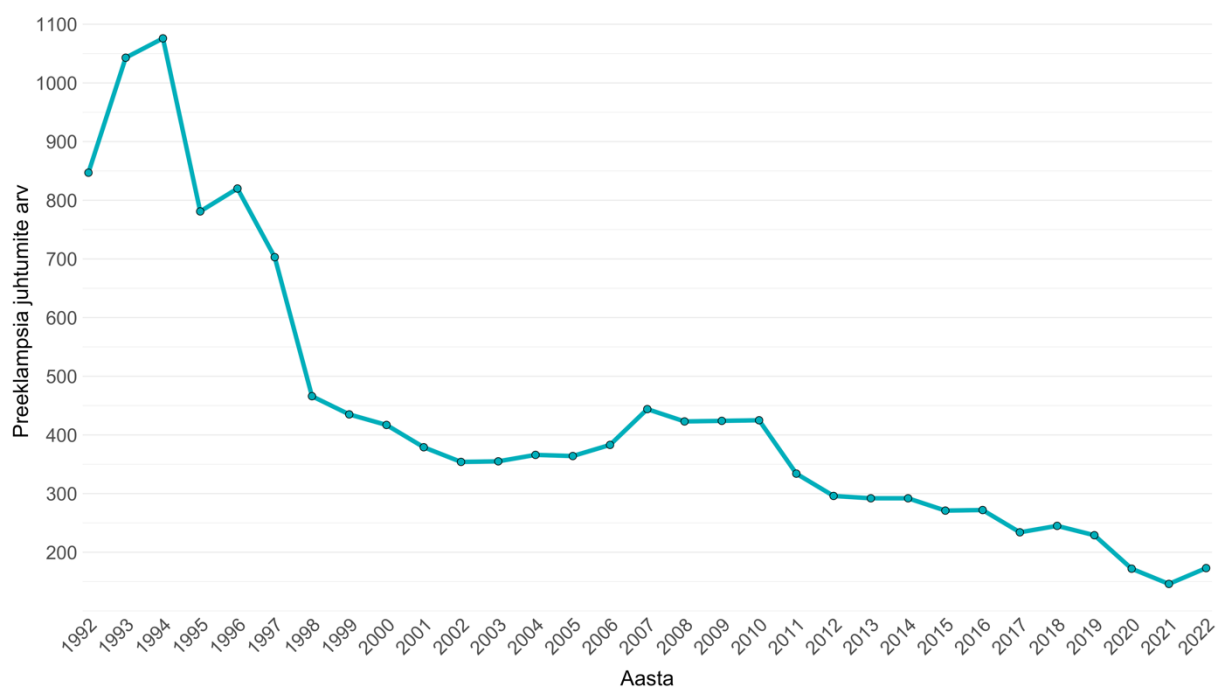
Nii preeklampsia kui rasedushüpertensiooni kujunemisel mängib rolli geneetiline eelsoodumus ning suuremahuliste ülegenoomsete seoseuuringute abil on juba ka tuvastatud mõned konkreetsed geneetilised riskilookused (Steinhorsdottir *et al.*, 2020; Honigberg *et al.*, 2023; Tyrmi *et al.*, 2023). Samas on täpsemalt kaardistamata preeklampsia ja rasedushüpertensiooni jagatud geneetika ning seosed teiste haiguste ja tunnustega.

Käesolevas töös uuriti preeklampsia ja rasedushüpertensiooni geneetikat Tartu Ülikooli Eesti Geenivaramu ja FinnGen andmete põhjal. Eksperimentaalse osa eesmärgiks oli analüüsida preeklampsia ja rasedushüpertensiooni tekke geneetilisi riskifaktoreid, hinnata nende omavahelist jagatud geneetikat ning geneetilisi seoseid ka teiste haiguste ning tunnustega. Antud magistritöö on valmistatud Tartu Ülikooli Molekulaar- ja rakubioloogia instituudi molekulaarsete bioteaduste õppekava raames. Töö eksperimentaalsena on teostatud TÜ Genoomika instituudis.

1. Kirjanduse ülevaade

1.1. Preeklampsia

Preeklampsia (PE) on hüpertensiivne ehk vererõhu tõusuga seotud raseduskomplikatsioon, mis võib ohustada nii lapse kui ka ema elu. Rahvusvahelise Haiguste Klassifikatsiooni¹ 10. väljaanne (RHK-10) järgi tähistatakse preeklampsiat tähtnumbrilise koodiga O14. Aastatel 2002–2010 diagnoositi preeklampsiat 4,6%-l kõigist rasedustest, kuid PE levimus varieerub piirkonniti ning esineb sagedamini Aafrika riikides (Lõuna-Aafrika Vabariik ja Nigeeria) ning Argentiinas (Abalos *et al.*, 2013; Abalos *et al.*, 2014). Eestis on viimase kolmekümne aasta jooksul preeklampsia juhtumite arv kahanenud (vt Joonis 1).



Joonis 1. Preeklampsia juhtumite arv Eestis aastatel 1992–2022. Andmed pärinevad Tervise Arengu Instituudi tervisestatistika ja terviseuuringute andmebaasist².

Ameerika Sünnitusabi ja Günekoloogia Kolledži (ACOG) poolt 2013. aastal välja antud suuniste kohaselt diagnoositakse preeklampsiat kui varasemalt normotensiivsel ehk normaalse vererõhuga naisel esineb pärast 20. rasedusnädalat püsiv kõrge vererõhk (süstoolne vererõhk \geq 140 mmHg ja/või diastoolne vererõhk \geq 90 mmHg), millega kaasneb esmakordne proteiinuuria ehk seisund, mille korral uriinis leidub valku (300 mg ööpäevases uriinis), või vähemalt üks järgnevatest sümptomitest: trombotsütopeenia ehk vereliistakute vähesus, neeru- või

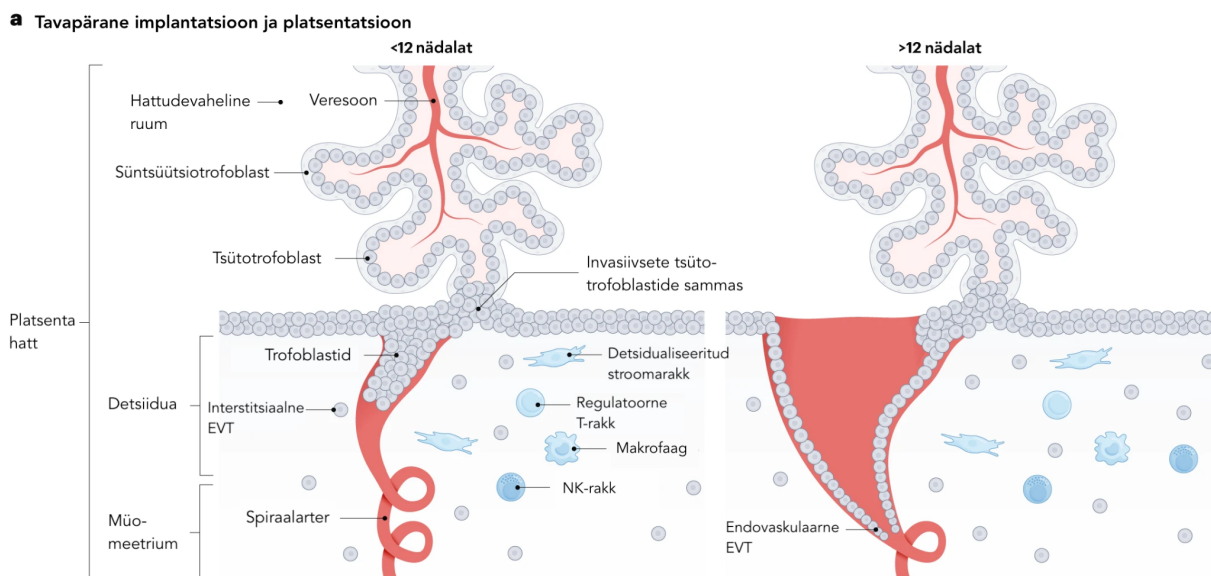
¹ <https://icd.who.int/browse10/2019/en>

² <https://statistika.tai.ee/>

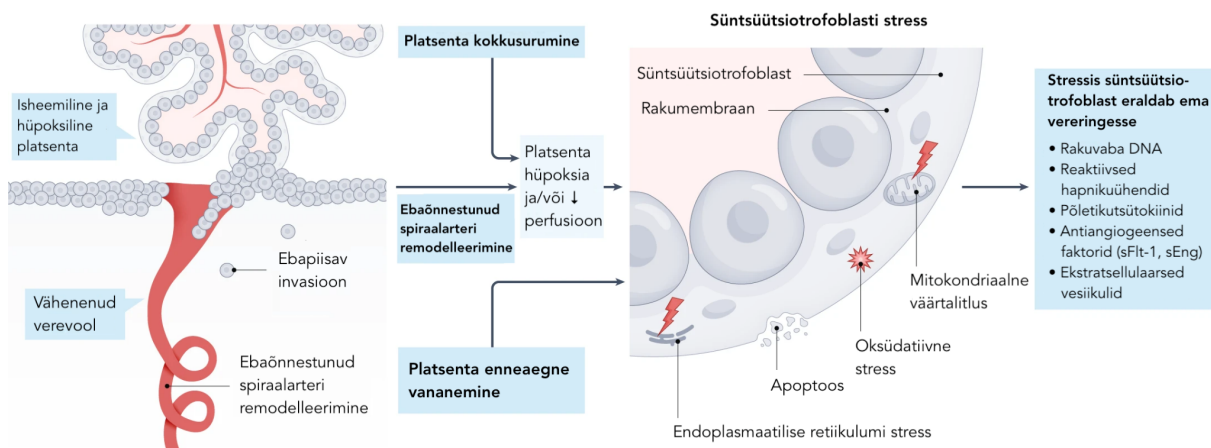
maksapuudulikkus, kopsuturse või esmakordne nägemishäire (American College of Obstetricians and Gynecologists, 2013).

On leitud mitmeid PE riskifaktoreid. Nende hulka kuuluvad näiteks hüpertensioon, krooniline neerupuudulikkus, esimene rasedus, preeklampsia või rasedushüpertensioon varasema raseduse ajal, suhkurtõbi, mitmikrasedus, viljatusravi, kehamassiindeks ($\geq 30 \text{ kg/m}^2$), vanus (≥ 35 a) ja PE esinemine sugulastel (ülevaade Mol *et al.*, 2016; Shen *et al.*, 2017; ülevaade Dimitriadis *et al.*, 2023). Viimane lubab eeldada, et preeklampsia kujunemisel mängib rolli ka geneetiline eelsoodumus. On viidud läbi mitmeid uuringuid, mis näitavad, et Euroopa päritoluga isikutel jääb preeklampsia päritavus 31–38% vahemikku (Nilsson *et al.*, 2004; Steinhorsdottir *et al.*, 2020). Kuna PE on oma olemuselt klassikaline komplekstunne (seisundi tekkel on oma roll nii geneetikal kui keskkonnal), on PE geneetilisi riskifaktoreid analüüsitud erinevates ülegenoomsetes seoseuuringutes, mis sobivad selliste tunnuste geneetika uurimiseks kõige paremini. Kokkuvõtte varasematest uuringutest on esitatud peatükis 1.4. Preeklampsia ja rasedushüpertensiooni senised ülegenoomsed seoseuuringud.

Haiguse tekkepõhjuste vaatevinklist on preeklampsia heterogeenne rasedustüsistus. Sõltuvalt preeklampsia kliiniliste sümptomite ilmnemise ajast, eristatakse PE-l kahte alatüüpi: varajane (avaldub enne 34. rasedusnädalat) ja hiline (avaldub alates 34. rasedusnädalast) (Lisonkova ja Joseph, 2013; Tranquilli *et al.*, 2013). On pakutud, et nii varajane kui ka hiline PE on tingitud süntsüütsiotrofoblasti stressist (I faas), kuid stressi põhjustaja ning algushetk on alatüüpide vahel erinevad. Stressis süntsüütsiotrofoblastid eritavad ema vereringesse antiangiogeenseid faktoreid nagu sFlt-1 (FMS-sarnase türosiini kinaas-1 lahustuv retseptor, *soluble fms-like tyrosine kinase-1*) ja sEng (lahustuv endogliin, *soluble endoglin*), mis kahjustavad ema vaskulaarendoteeli ja viivad PE kliiniliste sümptomite avaldumiseni (II faas) (ülevaade Staff, 2019; ülevaade Dimitriadis *et al.*, 2023). Varajase PE puhul peetakse süntsüütsiotrofoblastide stressi põhjustajaks häireid implantatsioonil (vt Joonis 2b). Ebaõnnestunud hattudeväliste trofoblastide invasioon ja ema spiraalarterite remodelleerimise tagajärjel ei saa platsenta piisavalt verd ja hapnikku. Hilise PE tekkepõhjusteks on välja pakutud koorioni hattude tihe paiknemine (platsenta kokkusurumine) ja platsenta enneaegne vananemine (vt Joonis 2b), mille tagajärjel platsenta varustamine hapniku ja toitainetega on samuti piiratud (ülevaade Dimitriadis *et al.*, 2023). Gyselaers'i ülevaade hemodünaamika varieeruvusest hilise PE siseselt toonitab, et hiline preeklampsia võib omakorda jaguneda kaheks alamtüübiks (ülevaade Gyselaers, 2022). Preeklampsia täpsed molekulaarsed tekkemehanismid on teadmata.



b Preeklampsiaga seotud süntsüüsiotrofoblasti stress



Joonis 2. (a) Vasakul pool on kujutatud tavapärase implantatsiooni ja platsentatsiooni ning paremal pool on remodelleeritud spiraalarter; **(b)** vasakul pool on kujutatud ebaõnnestunud spiraalarterite remodelleerimine ning paremal pool ebaõnnestunud spiraalarterite remodelleerimisest, platsenta kokkusurumisest ja platsenta enneaegsest vanemisest tekitatud süntsüüsiotrofoblasti stress (ülevaade Dimitriadis *et al.*, 2023, kohandatud). EVT – hatuväline trofoblast, sEng – lahustuv endogliin, sFlt-1 – FMS-sarnase türosiini kinaasi-1 lahustuv retseptor.

Preeklampsia üheks raskeimaks ilminguks on HELLP-sündroom (*hemolysis, elevated liver enzymes and low platelets*), mida iseloomustavad hemolüüs, suurenenud maksaensüümide tase ja madal vereliistakute arv. Eklampsia on harva esinev, kuid tõsine preeklampsia tüsistus, mille puhul lisanduvad klassikalistele PE sümptomitele krampid. Eklampsia tekke vältimiseks manustatakse ägedate preeklampsia sümptomitega rasedale vahetult enne ja pärast sünnitust magneesiumsulfaati (American College of Obstetricians and Gynecologists, 2013).

Ainus teadaolev preeklampsia ravi on sünnitus, mille järgselt kaovad PE sümptomid. Lisaks emade tervist ohustavatele riskidele tuleb sünnituse ajastamisel kaaluda ka lapse tervist, kuna

lühemast gestatsioonist tulenevalt esinevad enneaegsetel lastel mitmesugused komplikatsioonid (ülevaade Mol *et al.*, 2016). Üheks tõhusaks ja odavaks viisiks PE tüsistuste ärahoidmiseks on hiljemalt alates 16. rasedusnädalast riskigruppi kuuluvate naiste poolt aspiriini võtmine (Roberge *et al.*, 2013; Huai *et al.*, 2021). Sellegipoolest pole praeguseni kindlat ravimeetodit välja kujunenud ning on tõendeid, et aspiriini mõju ulatus oleneb etnilisest päritolust (Tolcher *et al.*, 2020).

Lisaks preeklampsia lühiajalistele mõjudele ema ja lapse tervisele, on leitud tõendeid tüsistuse negatiivsete tervisemõjude kohta ka hilisemas elus. Preeklampsiaga põdenud naitele keskendunud Wu *et al.* metaanalüüsist selgus, et PE kahekordistab teist tüüpi suhkurtõve väljakujunemise riski. Antud seos nõuab täiendavat uurimist, kuna preeklampsia ja teist tüüpi suhkurtõbi jagavad mitmeid riskifaktoreid (Wu *et al.*, 2016). Lisaks on leitud, et preeklampsiat põdenud naistel on suurem tõenäosus haigestuda südame-veresoonkonna haigustesse kui tervet rasedust kandnud naistel. Näiteks, kahekordistub insulti, südamepuudulikkuse ja koronaarterite haiguse risk (Wu *et al.*, 2017).

Eesti vastsündinute kaalu uurinud teadustöös ilmnis seos PE ja keskmisest madalama sünnikaalu vahel ning eklampsia negatiivne efekt sünnikaalule on veelgi suurem (Sildver *et al.*, 2015). Antud leid pakub huvi, sest madal sünnikaal omakorda suurendab täiskasvanueas kõrgvererõhutõve tekkimise riski (Chen *et al.*, 2012). Samuti on leitud seos preeklampsia ja järglaste kõrgeenenud süstoolse ning diastoolse vererõhu vahel ja ka kergelt suurenenud kehamassiindeksiga, mis on südame-veresoonkonnahaiguste riskifaktoriteks. Sealjuures on varajase PE järglaste süstoolne vererõhk süstemaatiliselt kõrgem kui hilise PE järglastel (Andraweera ja Lassi, 2019; Renlund *et al.*, 2023). Taani populatsioonipõhisest uuringust selgus, et preeklampsia ja eklampsiaga diagnoositud emade lastel esineb sagedamini insulti, peaaegu veresoonte haiguseid ja hüpertensiivsed häireid juba nooremas eas (enne neljakümnendat eluaastat) (Huang *et al.*, 2021).

1.2. Rasedusaegne hüpertensioon

Rasedushüpertensioon (GH) on teine peamine hüpertensiivne rasedustüsistus. Rasedushüpertensiooni iseloomustab uustekkeline kõrge vererõhk (süstoolne vererõhk ≥ 140 mmHg ja/või diastoolne vererõhk ≥ 90 mmHg), mis ilmub pärast 20. rasedusnädalat (American College of Obstetricians and Gynecologists, 2020). Viimasel ajal on hakanud ilmuma suuniseid, mille kohaselt klassifitseeritakse vererõhk $\geq 130/80$ mmHg 1. staadiumi hüpertensiooniks (ülevaade Ying *et al.*, 2018). See tähendab, et potentsiaalselt võiks ka raseduhüpertensiooni mõistet laiendada ning pakkuda sümptomite halvenemist ennetavaid

meetmeid rohkematele naistele. Rasedushüpertensiooni esinemissagedus varieerub piirkonniti, kuid jääb vahemikku 1.8–4.4% kõikidest rasedustest (ülevaade Umesawa ja Kobashi, 2017). RHK-10 klassifikatsiooni järgi on rasedushüpertensioonile omistatud tähtnumbriline kood O13¹.

Hiljutine Espeche jt uuring tõstatab tõsise valediagnoosimise probleemi. Vererõhu 24-tunni monitooringu abil on leitud, et rohkem kui pooled rasedusaegse hüpertensiooniga diagnoositud naised on põdenud maskeeritud hüpertensiooni (vererõhk \geq 130/80 mmHg) enne 20. rasedusnädalat. Maskeeritud hüpertensiooniga rasedate seas on PE tekkimise risk neli korda kõrgem kui tõelise rasedusaegse hüpertensiooniga naistel (Espeche *et al.*, 2022).

Rasedusaegse kõrgvererõhutõve riskifaktoriteks on esimene rasedus, mitmikrasedus, PE varasema raseduse ajal, suhkurtõbi ja kehamassiindeks (\geq 30 kg/m²) (Shen *et al.*, 2017). Rootsi suguvõsauuringus leiti, et rasedushüpertensiooni päritavus on 20% (Nilsson *et al.*, 2004). Rasedushüpertensiooni peetakse preeklampsia riskifaktoriks ning varasemalt on näidatud, et kuni neljandikul kõigist GH-ga diagnoositud naistest ilmneb preeklampsia (Saudan *et al.*, 1998). Ülevaade rasedushüpertensiooni geneetilistest riskifaktoritest on esitatud peatükis 1.4. Preeklampsia ja rasedushüpertensiooni senised ülegenoomsed seoseuuringud.

Rasedushüpertensiooniga kaasnevad PE-ga võrreldes rasked tüsistused harvem. Näiteks esineb enneaegset sünnitust kolm korda vähem (Shen *et al.*, 2017). Sellegipoolest on GH suureks koormuseks tervishoiusüsteemile. On teada, et rasedusaegne kõrgvererõhutõbi tõstab südameveresoonkonna haiguste riski hilisemas elus (Veerbeek *et al.*, 2015).

1.3. Ülegenoomsed assotsiatsiooniuringud

Peaaegu 20 aastat tagasi kasutusele võetud ülegenoomsed assotsiatsiooniuringud ehk GWAS-id (*genome-wide association study*) on muutunud oluliseks tööriistaks komplekstunnuste ja -haiguste tekkemehhanismide lahtimõtestamisel (ülevaade Uffelmann *et al.*, 2021). Populatsioonipõhiste biopankade abil järjest suurenevad valimid on võimaldanud saada väärtuslikku infot paljude komplekstunnuste ja -haiguste geneetilise tausta kohta. Nii näiteks uuriti hiljutises pikkuse GWAS analüüsis 5 miljoni inimese andmeid, mis võimaldas tuvastada sisuliselt kõik pikkusega seotud sagedased geenivariandid (Yengo *et al.*, 2022). Ülegenoomsete seoseuuringute meetod põhineb geneetiliste variantide (peamiselt ühenukleotiidsete polümorfismide ehk SNPde) ja tunnuste vahel seoste leidmisel. GWAS-i on võimalik viia läbi kasutades kogugenoomi või eksoomi sekveneerimise andmeid, kuid kiibipõhise

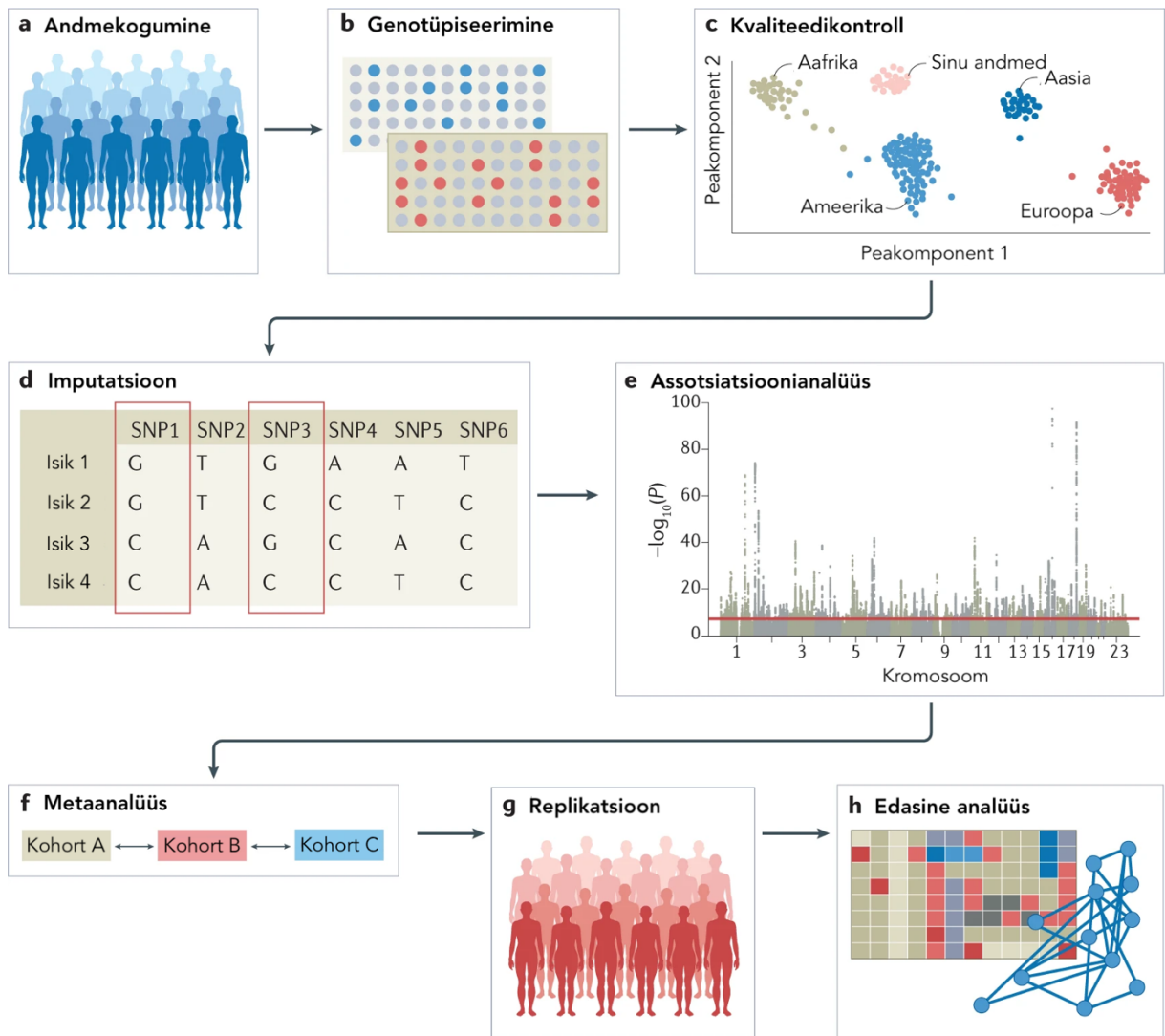
genotüpiseerimise odavuse tõttu kasutatakse assotsiatsiooniuuringutes peamiselt just genotüpiseerimiskiipidelt pärinevaid andmeid (ülevaade Tam *et al.*, 2019).

Ülegenoomsetel assotsiatsiooniuuringutel on suur potentsiaal (personaal)meditsiini edasiarendamisel, kuna GWAS-i tulemusi saab kasutada sisendina nii haigestumise riski ennustamisel, haiguste alatüüpide eristamisel, ravimite väljaarendamisel ja ka ravimite eluohtlike kõrvaltoimete prognoosimisel (ülevaade Manolio, 2013).

Riskigrupi kuuluvate indiviidide varajane tuvastamine ning haiguste ennetamine võimaldab edendada inimeste heaolu ja säästa ressursse (Kiflen *et al.*, 2022). Hiljutine Khera *et al.* uuring on näidanud, et GWAS-i tulemuste põhjal koostatud polügeensete skooride rakendamine viie komplekshaiguse puhul võimaldas tuvastada märkimisväärset osa rohkem kui kolmekordse haigestumise riskiga Ühendkuningriigi Biopanga populatsioonist (1,5%–8%) (Khera *et al.*, 2018). Raskesti eristatavate alatüüpidega haiguste puhul on assotsiatsiooniuuringu tulemused samuti osutunud kasulikuks.

Ülegenoomsete seoseuuringute töövoog koosneb mitmest etapist (vt Joonis 3). Esiteks, uuringusse kaasatakse isikuid, kellelt kogutakse geeniproov ja andmed fenotüübi kohta, näiteks küsitluse või digitaalse terviselugu kaudu. Teiseks, uuringu osalejate DNA genotüpiseeritakse kasutades mikrokiipe eelnevalt valitud markeritega. Kogutud andmed läbivad kvaliteedikontrolli ning genotüpiseerimata jäänud lüngad täidetakse imputatsiooni abil. Kolmandaks, viiakse läbi andmete statistiline analüüs tuvastamaks tunnusega assotsieerunud geneetilised variandid. Mittekohustuslikuks etapiks on metaanalüüs, mis võimaldab suurendada uuringu võimsuse. Neljandaks, saadud tulemusi tuleb kinnitada geneetiliselt sarnases sõltumatus kohordis. Uuringu väljundiks on tunnusega seotud geneetilised variandid ning nende efekti suurus. Saadud tulemusi interpreteeritakse selgitades, mis geenid ning bioloogilised protsessid on mõjutatud geneetiliste variantide poolt (ülevaade Uffelmann *et al.*, 2021).

GWAS analüüsi tulemusena saadud koondstatistikud on heaks sisendiks edasisteks analüüsideks, mis võimaldab paremini mõista huvipakkuvate tunnuste bioloogilist tausta. Siin on heaks näiteks nn geneetiliste korrelatsioonide analüüs. Geneetiline korrelatsioon näitab, millisel määral on kahe tunnuse geneetiline taust omavahel seotud. Selliste seoste analüüs võib anda infot näiteks haiguse tekkemehhanismide või riskifaktorite kohta. Üheks enimkasutatud meetodiks on nn LD-skoori regressioon, mis kasutab sisendina GWAS koondstatistikuid ning seega ei vaja indiviidi-tasemel andmeid (Bulik-Sullivan *et al.*, 2015).



Joonis 3. Skemaatiline ülevaade ülegenoomse assotsiatsiooniuringu töövoost (ülevaade Uffelmann *et al.*, 2021, kohandatud).

1.4. Preeklampsia ja rasedushüpertensiooni senised ülegenoomsed seoseuuringud

Senini teostatud suuremad preeklampsia ja rasedushüpertensiooni ülegenoomsed seoseuuringud on kokkuvõtvalt esitatud Tabelis 1. Väljatoodud assotsiatsiooniuuringute ülesehitus varieerub. On analüüsitud seoseid haigustega nii ema, lapse kui isa genoomi uuringutes. Assotsiatsioone isa genoomis on vaadeldud ainult Honigberg *et al.* uuringus, kuid ühtegi ülegenoomselt olulise geneetilise variandi pole tuvastatud. Tähelepanuväärne on asjaolu, et analüüsitud fenotüübid samuti varieeruvad uuringute vahel. On vaadeldud nii nn puhtaid diagnoose (preeklampsia, rasedushüpertensioon) kui mitmeid kombinatsioone diagnoosidest (preeklampsia või eklampsia (või eelnev hüpertensiivne häire lisandunud preeklampsiaga), preeklampsia või (rasedus)hüpertensioon).

Sarnaselt teiste komplekshaigustega on järjest suurenevad valimimahud võimaldanud tuvastada järjest rohkem seotud lookuseid, kusjuures korduvalt (> 2 korda) on näidatud *MTHFR*, *MECOM*, *FGF5*, *RGL3*, *SH2B3*, *ZNF831*, *FTO* ning *FLT1* seotust. *FLT1* lookuse seotust on näidatud nii ema kui lapse genoomi uuringud.

Tabel 1. Kokkuvõtte varasemate ülegenoomsete preeklampsia ja rasedushüpertensiooni seoseuringute tulemustest. EA – efekialleel (*effect allele*); AA – alternatiivne alleel (*alternative allele*); OR – šansside suhe (*odds ratio*); 95% CI – šansside suhte 95%-usaldusintervall (*confidence interval*). Mustas paksus kirjas on geneetilised variandid ja geenid, mida toodi välja kahes uuringus; oranžis paksus kirjas on geneetilised variandid ja geenid, mida toodi välja rohkem kui kahes uuringus.

Diagnoos	Genoom	Marker (EA/AA)	Geen	OR (95% CI)	p-väärtus	Juhud/ kontrollid	Viide
Preeklampsia/ eklampsia	Ema	rs149764880 (G/T)	<i>MTHFR</i> , <i>CLCN6</i>	1,11	$1,8 \times 10^{-8}$	17 150/ 451 241	(Honigberg <i>et al.</i> , 2023)
		rs708119 (C/G)	<i>WNT3A</i>	1,08	$2,3 \times 10^{-8}$		
		rs9855086 (A/T)	<i>MECOM</i>	1,08	$5,7 \times 10^{-10}$		
		rs16998073 (T/A)	<i>FGF5</i>	1,11	$3,1 \times 10^{-14}$		
		rs2442752 (C/T)	<i>MICA</i>	1,08	$2,3 \times 10^{-9}$		
		rs5899121 (TG/T)	<i>LINC00484</i>	1,12	$4,6 \times 10^{-8}$		
		rs2508372 (A/G)	<i>PGR</i>	1,10	$1,3 \times 10^{-9}$		
		rs10774624 (G/A)	<i>SH2B3</i>	1,10	$2,9 \times 10^{-12}$		
		rs7318880 (T/C)	<i>FLT1</i>	1,09	$6,7 \times 10^{-12}$		
		rs1421085 (C/T)	<i>FTO</i>	1,10	$4,7 \times 10^{-12}$		
		rs167479 (G/T)	<i>RGL3</i>	1,10	$7,6 \times 10^{-10}$		
		rs259983 (C/A)	<i>ZNF831</i>	1,13	$1,0 \times 10^{-12}$		
		rs2653414 (A/C)	<i>FGL1</i>	1,27	$1,0 \times 10^{-6}$		
		rs17572606 (T/C)	<i>UPBI</i>	1,45	$1,9 \times 10^{-7}$		
Rasedushüper- tensioon	Ema	rs9855086 (A/T)	<i>MECOM</i>	1,10	$4,6 \times 10^{-8}$	8 961/ 184 925	(Honigberg <i>et al.</i> , 2023)
		rs16998073 (T/A)	<i>FGF5</i>	1,12	$1,0 \times 10^{-10}$		
		rs13154066 (C/T)	<i>NPR3</i>	1,11	$5,2 \times 10^{-10}$		
		rs7139122 (A/G)	<i>TNS2</i> , <i>CSAD</i>	1,49	$2,4 \times 10^{-10}$		
		rs167479 (G/T)	<i>RGL3</i>	1,11	$1,5 \times 10^{-8}$		
		rs2208589 (G/A)	<i>PREX1</i>	1,15	$2,8 \times 10^{-8}$		
rs260017 (A/G)	<i>ZNF831</i>	1,17	$2,6 \times 10^{-12}$				

		rs17367504 (A/G)	<i>MTHFR, CLCN6</i>	1,13	$1,7 \times 10^{-7}$		
Preeklampsia/ eklampsia/ eelnev hüpertensiivne häire lisandunud preeklampsiaga	Ema	rs4245909 (G/A)	<i>MECOM</i>	0,92 (0,90–0,95)	$3,19 \times 10^{-9}$	16 743/ 271 306	(Tyrmi <i>et al.</i> , 2023)
		rs16998073 (T/A)	<i>FGF5</i>	1,12 (1,09–1,15)	$1,33 \times 10^{-15}$		
		rs2596471 (G/A)	<i>HLA, PSORS1C2</i>	1,11 (1,07–1,15)	$1,98 \times 10^{-9}$		
		rs7862828 (C/A)	<i>AUH, LINC00484</i>	1,10 (1,06–1,13)	$1,12 \times 10^{-8}$		
		rs3018700 (C/T)	<i>PGR, TRPC6</i>	1,13 (1,09–1,18)	$9,98 \times 10^{-10}$		
		rs10774624 (A/G)	<i>ATXN2, SH2B3</i>	0,92 (0,89–0,94)	$2,52 \times 10^{-10}$		
		rs7318880 (T/C)	<i>FLT1</i>	1,10 (1,07–1,13)	$5,04 \times 10^{-12}$		
		rs1421085 (C/T)	<i>FTO</i>	1,10 (1,07–1,13)	$1,55 \times 10^{-11}$		
		rs6026744 (T/A)	<i>ZNF831</i>	1,13 (1,09–1,17)	$9,72 \times 10^{-11}$		
Preeklampsia/ (rasedus)hüper- tensioon	Ema	rs13306561 (G/A)	<i>MTHFR, NPPA</i>	0,88 (0,84–0,91)	$7,49 \times 10^{-12}$	15 200/ 115 007	(Tyrmi <i>et al.</i> , 2023)
		rs1918969 (C/T)	<i>MECOM</i>	0,93 (0,90–0,95)	$1,34 \times 10^{-8}$		
		rs16998073 (T/A)	<i>FGF5</i>	1,14 (1,11–1,17)	$6,27 \times 10^{-19}$		
		rs12656497 (C/T)	<i>NPR3</i>	1,10 (1,07–1,13)	$2,52 \times 10^{-12}$		
		rs10882398 (A/T)	<i>PLCE1</i>	1,11 (1,08–1,14)	$1,77 \times 10^{-13}$		
		rs10843404 (C/G)	<i>PZP</i>	1,08 (1,05–1,11)	$3,16 \times 10^{-8}$		
		rs117928258 (T/C)	<i>TNS2</i>	1,45 (1,32–1,60)	$2,74 \times 10^{-14}$		
		rs6224 (T/G)	<i>FURIN, FES</i>	1,09 (1,06–1,12)	$1,77 \times 10^{-10}$		
		rs113935429 (A/AT)	<i>FTO</i>	1,10 (1,07–1,13)	$1,07 \times 10^{-10}$		
		rs167479 (T/G)	<i>RGL3</i>	0,90 (0,88–0,93)	$6,95 \times 10^{-13}$		
		rs979971 (T/C)	<i>ACTN4</i>	0,92 (0,90–0,95)	$9,10 \times 10^{-9}$		
		rs2208589 (G/A)	<i>PREX1</i>	1,12 (1,08–1,17)	$1,70 \times 10^{-8}$		
		rs201454025 (G/GTGT)	<i>ZNF831</i>	1,14 (1,10–1,18)	$8,41 \times 10^{-13}$		
Preeklampsia	Ema	rs1918975 (T/C)	<i>MECOM</i>	1,12 (1,07–1,16)	$2,5 \times 10^{-8}$	9515/ 157 719	(Steinthorsdottir <i>et al.</i> , 2020)
		rs1458038 (T/C)	<i>FGF5</i>	1,12 (1,07–1,16)	$4,2 \times 10^{-8}$		
		rs10774624 (G/A)	<i>SH2B3</i>	1,12 (1,07–1,16)	$7,0 \times 10^{-8}$		

		rs1421085 (C/T)	<i>FTO</i>	1,13 (1,08–1,17)	$2,0 \times 10^{-9}$		
		rs259983 (C/A)	<i>ZNF831</i>	1,15 (1,09– 1,22)	$2,2 \times 10^{-7}$		
Preeklampsia	Laps	rs4769612 (C/T)	<i>FLT1</i>	1,17 (1,12–1,23)	$3,0 \times 10^{-11}$	6775/ 375 372	(Steinthorsdottir <i>et al.</i> , 2020)

2. Eksperimentaalosa

2.1. Töö eesmärgid

Käesoleva magistritöö eesmärgiks oli analüüsida rasedushüpertensiooni ja preeklampsia emapoolseid geneetilisi riskifaktoreid ning seoseid teiste fenotüüpidega, et paremini mõista uuritavate haiguste patogeneesi. Selle saavutamiseks oli töö alaesmärkideks:

1. anoteerida Eesti ja Soome andmete GWAS metaanalüüsi tulemused kasutades FUMA veebiplatvormi;
2. analüüsida rasedushüpertensiooni ja preeklampsia geneetilisi seoseid teiste tunnuste ja reproduktiivsete diagnoosidega kasutades LD skoori regressiooni meetodit.

2.2. Materjalid ja meetodika

2.2.1. Valim ja uuritavad tunnused

Juhtkontrolluuringu läbiviimiseks kasutati Tartu Ülikooli Eesti Geenivaramu (TÜ EGV) andmeid. TÜ EGV doonorite hulka kuuluvad täisealised (vanus ≥ 18) isikud, kes on vabatahtlikult allkirjastanud teavitatud nõusoleku vormi, loovutanud oma geeniproovi (veeniverd) ning täitnud laiaulatusliku küsimustiku. Geenidoonorite andmeid täiendatakse pidevalt üleriigilise Tervisekassa andmebaasi kaudu (Leitsalu *et al.*, 2015). Eesti Geenivaramuga on liitunud üle 200 000 Eesti elaniku, mis hõlmab umbes 20% täiskasvanud populatsioonist. Peaaegu 135 000 geenidoonorit on naissoost (Pujol-Gualdo *et al.*, 2022). Käesolevas töös esitatud analüüsid on teostatud Eesti bioetika ja inimuuringute nõukogu poolt väljastatud eetikaloa 1.1-12/624 raames ning kasutades andmeväljastust N05.

Uuringu rasedushüpertensiooni ja preeklampsia juhtudeks valiti naised, kes on diagnoositud RHK-10 järgi O13 või O14 koodiga. Enne metaanalüüsi teostamist, viidi läbi GWAS TÜ EGV andmetel. Ülegenoomselt olulisi (p -väärtus $< 5 \times 10^{-8}$) geneetilisi variante ei õnnestunud tuvastada. Seega on käesolevas magistritöös esitatud vaid metaanalüüsi tulemused. Kontrollide valimisel kasutati kahte lähenemist. Ühes analüüsis (Variant 1) võeti kontrollideks kõik vastava diagnoosita naisgeenidoonorid, mistõttu kuuluvad kontrollidesse ka naised, kes pole kunagi rasedad olnud. Teises analüüsis (Variant 2) võeti kontrollideks vastava diagnoosita naisgeenidoonorid, kes on olnud rasedad. Kuna tulemuste võrdlemisel olulisi erinevusi ei tuvastatud, kasutati statistilise võimsuse huvides edasistes metaanalüüsides Variant 1 tulemusi ning need on esitatud ka käesolevas magistritöös.

Rasedushüpertensiooni seoseuuringus kasutatud juhtude ja kontrollide keskmine vanus biopangaga liitumisel oli TÜ EGV kohordis vastavalt 45,0 (SD = 16,3) ja 34,3 (SD = 8,23) aastat. Preeklampsia assotsiatsiooni uuringus kasutatud juhtude ja kontrollide keskmine vanus biopangaga liitumisel oli TÜ EGV kohordis vastavalt 44,9 (SD = 16,3) ja 34,3 (SD = 8,14) aastat.

Allpool kirjeldatud metaanalüüsis kombineeriti TÜ Eesti Geenivaramu andmete põhjal saadud tulemused Soome FinnGen andmebaasi 7. väljaande (R7) seoseuuringute koondstatistikutega. FinnGen R7 kuulub üle 300 000 geenidonorit, kellest ligikaudu 175 000 on naissoost³. FinnGen andmebaas ei väljasta indiviidi tasandil andmeid, mistõttu kohordi juhtude ja kontrollide keskmist vanust ei saa välja arvutada. Sellegipoolest on teada, et FinnGen kohordi keskmine vanus on 63 aastat (Kurki *et al.*, 2023). Uuritavate diagnooside vasted FinnGen R7 andmebaasis on toodud välja Tabelis 2.

Töös kasutatud valimite suurused on toodud välja Tabelis 3.

Tabel 2. Uuritavate diagnooside vasted RHK-10 järgi ja FinnGen R7 andmebaasis.

Diagnoos	RHK-10 kood	FinnGen R7
Rasedushüpertensioon	O13	O15_GESTAT_HYPERT
Preeklampsia	O14	O15_PREECLAMPS

Tabel 3. Uuritavate diagnooside juhtude ja kontrollide arv TÜ Eesti Geenivaramu ja FinnGen R7 kohortides.

Diagnoos	TÜ EGV		FinnGen R7		Metaanalüüs	
	Juhud	Kontrollid	Juhud	Kontrollid	Juhud	Kontrollid
Rasedushüpertensioon	2509	124 298	6562	160 670	9071	284 968
Preeklampsia	1378	125 429	5265	160 670	6643	286 099

2.2.2. Genotüpiseerimine, kvaliteedikontroll, imputatsioon

Käesolevas magistritöös kasutati olemasolevaid genotüpiseerimise andmeid. Geenidonorite proovide genotüpiseerimine, andmete kvaliteedikontroll ja imputeerimine toimusid keskselt TÜ Eesti Geenivaramu tuumiklaboris.

³ https://www.finnngen.fi/en/access_results

Genotüpiseerimist teostati kasutades *Illumina*⁴ GSAv1.0, GSAv2.0 ja GSAv2.0_EST kiipe.

Genotüpiseerimise andmed läbisid kvaliteedikontrolli allpool ära toodud parameetrite järgi:

- edukalt genotüpiseeritud markerite protsentuaalne osakaal (*call rate*, CR);
- genotüübilise ja fenotüübilise soo mittevastavus;
- Hardy-Weinbergi tasakaalustatus (*Hardy-Weinberg equilibrium*, HWE);
- harvema alleeli sagedus (*minor allele frequency*, MAF).

Analüüsist jäeti välja isikuid, kelle genotüübilise ja fenotüübilise soo vahel esines ebakõla või kelle CR jäi alla 95%. Geneetilised variandid, mille CR < 95%, HWE testi p-väärtus < 10⁻⁴ (autosomaalsed variandid) ning MAF < 1%, jäeti samamoodi välja. Geneetiliste variantide positsioonid määrati inimese referents genoomi versioon 37 (GRCh37) järgi (Pujol-Gualdo *et al.*, 2022). Eelfaasimiseks kasutati Eagle v2.3 tarkvara, mis on efektiivne ka väiksemate kohortide puhul (Loh *et al.*, 2016; Pujol-Gualdo *et al.*, 2022). Genotüpiseerimata SNP-de imputeerimiseks kasutati Beagle v.28Sep18.793 tarkvara ning Eesti-spetsiifilist referentspaneeli, mis koosneb 2297 Eesti elaniku kogugenoomidest (Browning ja Browning, 2007; Mitt *et al.*, 2017; Pujol-Gualdo *et al.*, 2022).

Soome kohordi geeniproovide genotüpiseerimine, kvaliteedikontroll ja imupatsioon on kirjeldatud Kurki *et al.* (2023) teadusartiklis ja FinnGen dokumentatsioonis⁵. Genotüpiseerimiseks kasutati nii *Illumina* kui *Affymetrix* kiipe, kuid kiibid pole täpsustatud. Geneetiliste variantide positsioonid määrati GRCh38 järgi. Eelfaasimiseks kasutati Eagle v2.3.5 tarkvara (Loh *et al.*, 2016). Imputeerimiseks kasutati Beagle v.08Jun17.d8b tarkvara ning Soome-spetsiifilist referentspaneeli *Sequencing Initiative Suomi* (SISu) v.3⁶ (Browning ja Browning, 2007).

2.2.3. Assotsiatsioonianalüüs TÜ Eesti Geenivaramu ja FinnGen R7 andmetel

Assotsiatsioonianalüüs TÜ EGV andmetel viidi läbi kasutades tarkvarapaketti REGENIE v2.2.4. Meetodi eeliseks on lähenemine, mille kohaselt grupeeritakse uuritavaid SNP-e alaplökkideks. Tänu sellele on REGENIE kiire ja mälukasutuselt efektiivne ning võimaldab analüüsida mitu tunnust korraga. Lisaks võimaldab REGENIE segamudel jätta analüüsi sisse suguluses olevaid isikuid (Mbatchou *et al.*, 2021). Analüüsis kasutati kovariaatidena

⁴ <https://www.illumina.com/>

⁵ <http://r7.finnngen.fi>

⁶ <http://www.sisuproject.fi/>

sünniaastat ja 10 esimest geneetilist peakomponenti (PC). Madala imputatsiooni kvaliteediga (INFO *score* < 0,4) geneetilised variandid ning variandid, mille harvema alleeli koguarv < 5, jäeti analüüsist välja (Pujol-Gualdo *et al.*, 2022). Analüüsi teostas doktorant Natàlia Pujol Gualdo.

Nagu eelpool mainitud, Soome kohordi andmeid metaanalüüsi jaoks saadi seoseuuringute koondstatistikute kujul. Soome GWAS-i etapid on täpsemalt lahti kirjutatud FinnGen kodulehel⁷. Analüüsist jäeti välja variandid, mille imputatsiooni INFO *score* < 0,6 ja MAF < 0,0001. Erinevalt varasematest FinnGen väljaannetest, FinnGen R7-s esitatud koondstatistikud on saadud REGENIE v2.0.2 tarkvara abil (Mbatchou *et al.*, 2021). Kovariaatidena võeti arvesse (sugu), vanust, genotüüpiseerimise partiid ja 10 esimest geneetilist PC-d.

2.2.4. GWAS metaanalüüs TÜ Eesti Geenivaramu ja FinnGen 7R andmetel

Rasedushüpertensiooni ja preeklampsia GWAS metaanalüüsi teostas doktorant Natàlia Pujol Gualdo tarkvara GWAMA v2.1 abil (Mägi ja Morris, 2010). Ülevaade metaanalüüsi kaasatud valimist on toodud välja Tabelis 3 (vt peatükk 2.2.1. Valim ja uuritavad tunnused). Järgnevates analüüsides kasutati geneetilisi variante, mis esinesid nii TÜ EGV kui FinnGen R7 kohordis ning mille harvema alleeli esinemissagedus oli > 1%. Lõplikku analüüsi jäi seega 12 363 274 (rasedushüpertensioon) ja 12 363 270 (preeklampsia) geneetilist varianti.

2.2.5. GWAS metaanalüüsi tulemuste annoteerimine ja visualiseerimine FUMA veebiplatvormi abil

Ülegenoomse assotsiiooniuringu väljundiks on koondstatistikute fail, kus iga testitud SNP-i kohta on toodud välja üldine informatsioon (positsioon, efekialleel, alternatiivne alleel jms), beeta ehk efekti suurus ja suund ning assotsiatsioonitesti p-väärtus. Sageli võib assotsiatsiooniuuringust ilmnedu mitmeid statistiliselt olulisusi (p-väärtus < 5×10^{-8}) SNP-e, mis paiknevad geenidevahelises alas, mistõttu nende bioloogilise tähenduse lahtimõtestamine on raskendatud. Antud probleemi lahendamiseks on loodud funktsionaalse kaardistamise ja annoteerimise (FUMA) veebiplatvorm⁸, mis koondab endasse informatsiooni mitmest bioloogilisest andmekogust ning aitab kaardistada potentsiaalseid põhjuslike SNP-e ja gene (Watanabe *et al.*, 2017).

⁷ https://www.finnngen.fi/en/access_results

⁸ <https://fuma.ctglab.nl/>

2.2.5.1. Kandidaatvariantide ja riskilookuste kaardistamine

Kandidaat-SNP-de kaardistamiseks kasutati FUMA-s olevat funktsiooni SNP2GENE. Selleks laeti üles metaanalüüside koondstatistikute failid, milles iga analüüsitud variandi kohta toodi välja kromosoom, positsioon ja p-väärtus. FUMA automaatselt visualiseerib GWAS-i assotsiatsioonid Manhattani joonisele.

Juhtivate ja kandidaat SNP-de määramisel kasutati parameetrite vaikumisi kehtivaid väärtusi:

- juhtiva SNP-i p-väärtus $< 5 \times 10^{-8}$;
- oluliste sõltumatute SNP-de ahelduse tasakaalutuse (r^2) $\geq 0,6$;
- juhtivate SNP-de ahelduse tasakaalutuse (r^2) $\geq 0,1$;
- referentspaneel 1000G Phase3 EUR (Auton *et al.*, 2015);
- kahe LD-ploki grupeerimisel maksimaalne lubatud kaugus 250 kb.

Kandidaatvariantide prioritiseerimiseks tuvastatakse kõik olulised sõltumatud SNP-d, mille p-väärtus $< 5 \times 10^{-8}$ ja ahelduse tasakaalutus r^2 ühtegi teise variandiga ei ületa 0,6. Järgnevalt kaardistatakse iga olulise sõltumatu SNP-iga ahelduse tasakaalutuses ($r^2 \geq 0,6$) olevaid variante. Nii saadakse nn LD-plokke. Kõik lähestikku asetsevad (kaugus < 250 kb) LD-plokid grupeeritakse. Viimasena tuvastatakse LD-ploki juhtivaid SNP-e ehk SNP-e, mille r^2 ühtegi teise SNP-iga ei ületa 0,1. Tulemusena saadakse riskilookuseid, mis koosnevad ühest või mitmest juhtivast ning oluliselt sõltumatutest variantidest.

Kandidaatvariantide funktsionaalse rolli lahtimõtestamiseks koostab FUMA integreeritud informatsiooni põhjal kokkuvõttetabeleid. Näiteks integreeritud tööriist ANNOVAR kirjeldab kaardistatud variantide teadaolevat mõju geenidele (Wang *et al.*, 2010). Mittekodeerivate geneetiliste variantide funktsionaalseks annoteerimiseks on integreeritud ka CADD (*Combined Annotation–Dependent Depletion*, kombineeritud annotatsiooni-põhine kurnamine) v1.4 meetod, mis väljastab SNP-i C skoori vastavalt selle regulatoorsele efektile, patogeensusele, assotsioonidele komplekstunnustega jms (Kircher *et al.*, 2014). Nii on kõrge C skoor SNP-i kahjulikkuse indikaatoriks. Lisaks lingib FUMA kaardistatud SNP-de GWAS kataloogiga, mis aitab leida esile toodud variantide seoseid teiste tunnustega varasemate seoseuuringute tulemuste põhjal (Welter *et al.*, 2014).

2.2.5.2. Kandidaatgeenide kaardistamine

Ehkki FUMA annab ammendavat ülevaadet komplekstunnustega assotsieerunud geneetilistest variantidest, edasine kandidaatgeenide kaardistamine aitab paremini mõista, millised bioloogilised rajad on kaasatud haiguste väljakujunemisel.

Kandidaatgeenide kaardistamiseks on FUMA-sse integreeritud kolm erinevat võimalust:

1. MAGMA v1.08 geeni-põhine kaardistamine. Kaardistamine põhineb mitmesel lineaarsel regressioonimudelil, mida kasutatakse geenide p-väärtuste väljaarvutamiseks. Antud meetod võtab arvesse geenis paiknevate SNP-de assotsiatsioone uuritava tunnusega ja ahelduse tasakaalutust (Leeuw *et al.*, 2015). Vätiaks mitmesest testimisest (testitud valkude arv on 18 895) tingitud valepositiivseid tulemusi, statistiliselt olulisteks peeti vaid neid gene, mille p-väärtus $< 2,646 \times 10^{-6}$.
2. Geeniekspressiooni taset mõjutavatel lookustel (eQTL, *expression quantitative trait loci*) põhinev kaardistamine. Kaardistatakse gene, mille ekspressiooniga on seotud riskilookuse olulised sõltumatud SNP-d ning nendega LD-s olevad SNP-d (Watanabe *et al.*, 2017).
3. Kromatiininteraktsioonidel põhinev kaardistamine. Geenide identifitseerimiseks kasutatakse informatsiooni kandidaatvariantide kromatiininteraktsioonidest. Geeni kaardistatakse kui selle promootor kattub SNP-i interaktsioonipiirkonnaga (Watanabe *et al.*, 2017).

Lisaks arvestati kandidaatgeenide kaardistamisel juhtivate SNP-de lähimate geenidega. Kandidaatgeenide kaardistamisel FUMA-s kasutati parameetrite vaikimisi kehtivaid väärtusi ja andmestike, kuid eQTL-põhises geenide kaardistamises jäeti välja GTEx kataloogi varasemad versioonid (v6 ja v7). Nii eQTL kui kromatiininteraktsioonidel põhinev kaardistamine võimaldab kaardistada kandidaatgeene, mis paiknevad riskilookusest väljastpoolt. Nendel kahel meetoditel tuvastatud kandidaatgeenid kanti FUMA platvormil *circos*'e joonisele.

2.2.6. LD skoori regressioon

Diagnooside SNP-de abil seletatava päritavuse (h_{SNP}^2) ja tunnustevahelise jagatud geneetika (r_g) hindamiseks kasutati GWAS metaanalüüsi koondstatistikuid ja LD skoori regressiooni (LDSC, *linkage disequilibrium score regression*) meetodi v1.0.1⁹ (Bulik-Sullivan *et al.*, 2015 a; Bulik-Sullivan *et al.*, 2015 b). LD skoori regressiooni käigus filtreeritakse metaanalüüsi failid nii, et

⁹ <https://github.com/bulik/ldsc>

alles jäävad vaid sagedased variandid, mis kattuvad HapMap3 paneelis¹⁰ esitatud SNP-dega (kokku ligikaudu 1 400 000 SNP-d). LDSC regressiooni teostas doktorant Natàlia Pujol Gualdo.

2.2.6.1. Geneetilised korrelatsioonid

Geneetiliste korrelatsioonide analüüsis kasutati lisaks käesolevas magistritöös analüüsitud tunnustele ka teiste naisspetsiifiliste haiguste GWAS metaanalüüside koondstatistikuid, millega meie uurimisgrupp töötab (47 raseduse ja naiste tervisega seotud diagnoosi). Selline lähenemine võimaldab hinnata preeklampsia ja rasedushüpertensiooni seost ka teiste rasedusaegsete diagnoosidega ning naiste tervisega laiemalt. Liiga madala päritavushinnangu (mediaan -0,02435) ja või fenotüüpide kattuvuse tõttu eemaldati edasisest analüüsist 9 diagnoosi. Analüüsi tulemused visualiseeriti R v4.2.2 tarkvara¹¹ “*corrplot*” paketi abil. Tunnused järjestati hierarhilise klasterdamise (*order = “hclust”*) meetodi abil.

Komplekstunnuste geneetika virtuaallaborisse (CTG–VL, *Complex Traits Genetics Virtual Lab*¹²) integreeritud LDSC abil vaadeldi järgnevalt geneetilisi korrelatsioone PE ja rasedushüpertensiooni ning Ühendkuningriigi biopanga (UKBB) 1403 tunnuse vahel.

Tekstifaile andmetega laeti üles veebiplatvormile kujul:

CHR POS SNP A1 A2 EAF BETA SE P N,

kus CHR – kromosoom, POS – positsioon kromosoomil, SNP – SNP-i unikaalne kood, A1 – efektilleel, A2 – alternatiivne alleel, EAF – efektilleeli sagedus, BETA – efektilleeli efektsuurus, SE – efektsuuruse standardviga, P – p-väärtus, N – valimi suurus.

CTG–VL abil saadud tulemusi visualiseeriti R v4.2.2 tarkvara “*ggplot2*”, “*tidyverse*” ja “*forcats*” pakettide abil.

Mõlemas analüüsis saadud geneetiliste korrelatsioonide p-väärtustele teostati mitmese testimise korrektsiooni kasutades R v4.2.2 tarkvara „*p.adjust*“ funktsiooni „*fdr*“ meetodit ning tulemuste peatükis 2.3.3.2. Geneetilised korrelatsioonid tuuakse välja q-väärtused ehk korrigeeritud p-väärtused.

¹⁰ <https://www.broadinstitute.org/medical-and-population-genetics/hapmap-3>

¹¹ <https://www.r-project.org/>

¹² <https://vl.genoma.io/>

2.2.6.2. Päritavuse hinnang

LD skoori regressiooni üheks väljundiks on tunnuse päritavuse hinnang. Magistritöös uuritavad fenotüübid on juht-kontroll-tüüpi tunnused ning sellistel tunnustel puudub pidevatele tunnustele omane varieeruvus, mis võimaldaks arvutatud päritavushinnanguid arusaadavalt tõlgendada. Lisaks võib diagnoosi osakaal uuritavas kohordis erineda diagnoosi osakaaluga populatsioonis. Päritavuse hinnangu tõlgendamiseks ja valimi kallutatusega arvestamiseks on pakutud vaadeldud päritavuse ($h_{SNP O}^2$, *observed-scale heritability*) ja selle standardhälbe teisaldamist riskiskaala päritavusele ($h_{SNP L}^2$, *liability-scale heritability*) (Lee *et al.*, 2011). Päritavuskoeffitsientide teisaldamiseks kasutati R v4.2.2 tarkvaras Lee jt artiklist tuletatud valemit¹³:

$$h_{SNP L}^2 = h_{SNP O}^2 \times K^2 \times (1 - K^2) \div P \div (1 - P) \div z^2$$

K on tunnuse levimust populatsioonis, mida võrdsustati juhtude osakaaluga (P).

Uuritud 38 rasedusaegse ja naiste terviseiga seotud diagnoosi $h_{SNP L}^2$ visualiseeriti R v4.2.2 tarkvara “*ggplot2*”, “*tidyverse*” ja “*forcats*” pakettide abil.

¹³ <https://gist.github.com/nievergeltlab/fb8a20feded72030907a9b4e81d1c6ea>

2.3. Tulemused

2.3.1. Geneetiliste variantide prioritseerimine FUMA-s

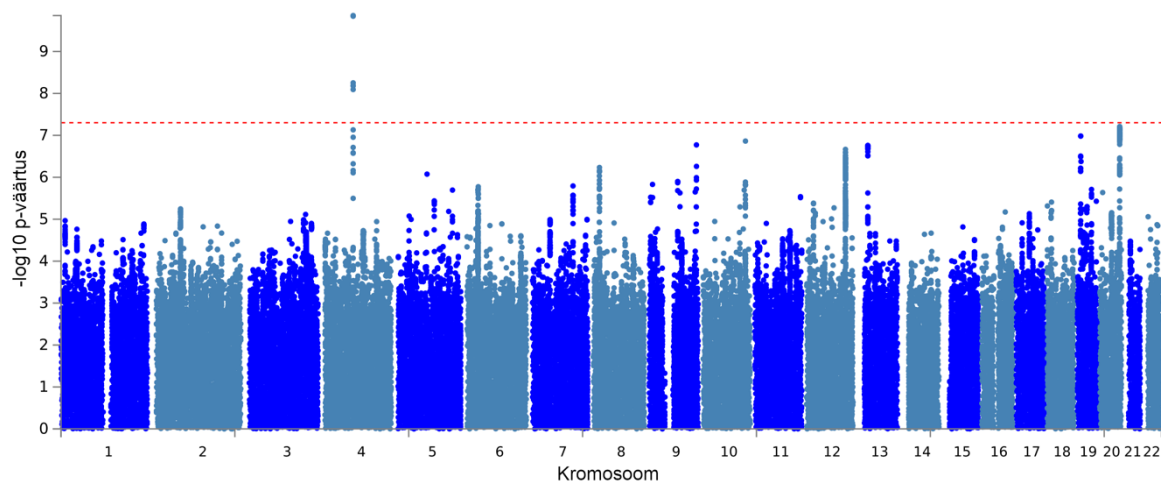
Selgitamaks välja, millised geenid võivad olla kaasatud preeklampsia ja rasedushüpertensiooni tekkemehhanismidesse, visualiseeriti ülegenoomsete assotsiatsiooni metaanalüüside tulemused FUMA platvormi SNP2GENE funktsiooni abil. Kõigepealt kaardistati üksikvariantide assotsiatsioone uuritavate diagnoosidega. Veebiplatvorm koostab GWAS-i tulemuste kokkuvõtte, kus on välja toodud tunnusega seotud geneetilised (riski)lookused ja nendes asuvad statistiliselt olulised (p -väärtus $< 5 \times 10^{-8}$) juhtivad geneetilised variandid. Ülevaade tuvastatud riskilookuste juhtivatest SNP-dest on esitatud Tabelis 4.

Preeklampsia GWAS metaanalüüsi tulemuste põhjal õnnestus kaardistada üks lookus 4. kromosoomis (vt Joonis 4). Antud lookuse juhtivaks SNP-iks on geneetiline variant rs16998073 (p -väärtus = $1,39 \times 10^{-10}$), mis paikneb ligikaudu 3 kb ülesvoolu *FGF5* geenist. Juhtiva SNP-i assotsiatsiooni preeklampsia ja ka rasedushüpertensiooniga on kirjeldatud varasemates haiguste GWAS-des (Honigberg *et al.*, 2023; Tyrmi *et al.*, 2023).

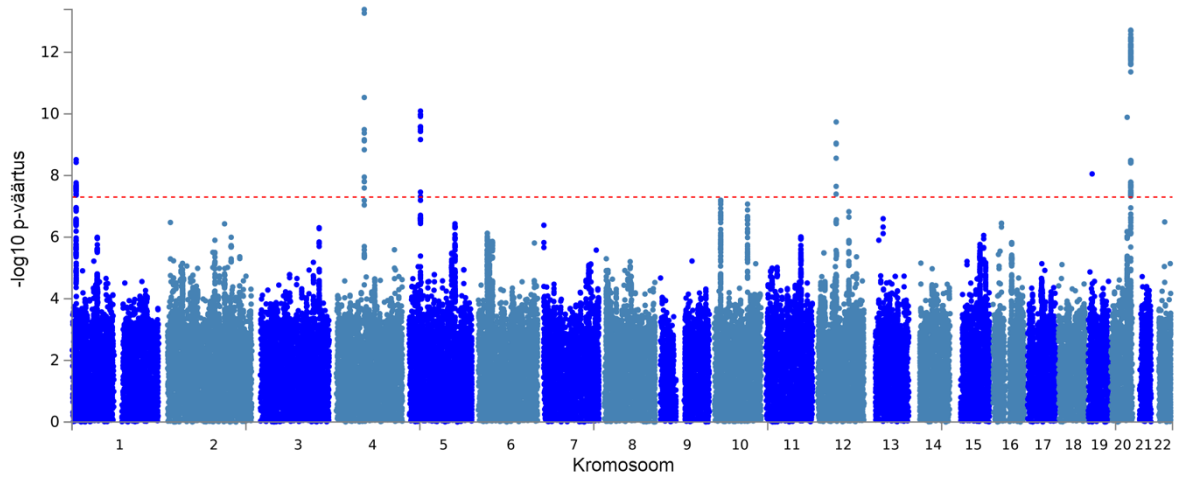
Rasedushüpertensiooni assotsiatsiooni metaanalüüsis leiti kuues kromosoomis (kromosoomid 1, 4, 5, 12, 19, 20) kokku seitse lookust (vt Joonis 5). Statistiliselt kõige olulisem geneetiline variant rs16998073 (p -väärtus = $4,08 \times 10^{-14}$) kattub PE uuringu ainsa statistiliselt olulise leiuga. Teiseks assotsieerunud variandiks on rs74715624 (p -väärtus = $1,91 \times 10^{-13}$), mis paikneb 20. kromosoomis. Variant asub ligikaudu 13 kb ülesvoolu *ZNF831* geenist. ANNOVAR-i järgi antud riskilookuse sõltumatu SNP-ga rs73618647 LD-s ($r^2 = 0,97$) olev variant rs56057707 (p -väärtus = $4,72 \times 10^{-8}$) paikneb *ZNF831* geeni eksonis ning on mittesünonüümne. Samas kromosoomis tuvastati ka teine statistiliselt oluline lookus. Riskilookuse juhtivaks SNP-iks on rs2208589 (p -väärtus = $1,29 \times 10^{-10}$), mis paikneb *PREX1* geeni intronis. Antud SNP (rs2208589) on näidanud seost ka varasemates rasedushüpertensiooni assotsiatsiooniuuringutes (Honigberg *et al.*, 2023; Tyrmi *et al.*, 2023). 1. kromosoomis paikneva lookuse juhtiv SNP on *KIAA2013* geeni introonne variant rs143439093 (p -väärtus $3,05 \times 10^{-9}$). Antud riskilookuse olulise sõltumatu SNP-ga rs13306561 LD-s ($r^2 = 0,62$) olev variant rs41275456 on mittesünonüümne SNP *C1orf167* geeni eksonis, mille C skoor = 15,28. ANNOVAR-i järgi on sama lookuse teise sõltumatu variandiga rs17376328 seotud ($r^2 = 0,99$) SNP rs55741089 (p -väärtus = $4,33 \times 10^{-8}$) *CLCN6* eksoni mittesünonüümne variant. Samuti tuvastati riskilookus 5. kromosoomis. Juhtiv variant rs12656497 (p -väärtus = $8,09 \times 10^{-11}$) paikneb kahest kattuvast geenist (*NPR3* ja *AC026703.1*) 40 kb allavoolu. Lisaks kaardistati riskilookust 12. kromosoomil. Lookuse juhtiv SNP rs7139122 (p -väärtus = $1,83 \times 10^{-10}$) paikneb *CSAD* geeni

intronis. Viimane tuvastatud juhtiv SNP rs167479 (p -väärtus = $8,85 \times 10^{-9}$) asub 19. kromosoomis, *RGL3* geeni eksonis, ning omab kõrget C skoori (24,4), mis viitab SNP-i negatiivsele mõjule valgu funktsioonile. ANNOVAR annotatsiooni kohaselt on tegemist mittesünonüümse SNP-ga.

Mõistmaks paremini uuritavate diagnooside jagatud geneetilist komponenti, kontrollisime rasedushüpertensiooni assotsiatsioonianalüüsis tuvastatud juhtivate variantide p -väärtused preeklampsia metaanalüüsis saadud assotsiatsioonidega (vt Lisa 1). Nagu eelpool mainitud, vaid üks geneetiline variant (rs16998073) näitas statistilist olulisust PE seoseuuringus. Sellegipoolest on tähelepanuväärsed kaks lisaassotsiatsiooni – rs167479 ja rs74715624 – mille p -väärtused jäid vahemikku $0,05 \times 10^{-6}$ – $0,05 \times 10^{-8}$ (*subsignificant*). Statistilise olulisuse piiri lähedal p -väärtused viitavad sellele, et antud geneetilised variandid võivad olla seotud ka PE tekkemehhanismidega.



Joonis 4. Preeklampsia seoseuuringute metaanalüüsi tulemuste Manhattan'i joonis. X-teljele on geneetiliste markerite skemaatilised asukohad genoomil. Y-teljel on vastava assotsiatsiooni p -väärtuse negatiivne kümnendlogaritm. Punane katkendjoon märgib mitmesele testimisele kohandatud olulisuse piiri (p -väärtus $\leq 5 \times 10^{-8}$). Joonis on tehtud FUMA veebiplatvormi abil.



Joonis 5. Rasedushüpertensiooni seoseuringute metaanalüüsi tulemuste Manhattan'i joonis. X-teljele on geneetiliste markerite asukohad genoomil. Y-teljel on vastava assotsiatsiooni p-väärtuse negatiivne kümnendlogaritm. Punane katkendjoon märgib mitmesele testimisele kohandatud olulisuse piiri ($p\text{-väärtus} \leq 5 \times 10^{-8}$). Joonis on tehtud FUMA veebiplatvormi abil.

Tabel 4. Preeklampsia ja rasedushüpertensiooni assotsiatsiooni metaanalüüside tulemused. EAF – efekti alleeli sagedus (*effect allele frequency*); OR – šansside suhe (*odds ratio*); 95% CI – šansside suhte 95%-usaldusintervall (*confidence interval*).

Diagnoos	Marker	Kromosoom:positsioon	Efektialleel/ teine alleel	EAF	OR (95% CI)	p-väärtus	Lähim geen
Preeklampsia	rs16998073	4:81184341	A/T	0,67	0,88 (0,85–0,92)	$1,39 \times 10^{-10}$	<i>FGF5</i>
	rs143439093	1:11981282	A/G	0,94	1,22 (1,14–1,30)	$3,05 \times 10^{-9}$	<i>KIAA2013</i>
	rs16998073	4:81184341	A/T	0,67	0,88 (0,85–0,91)	$4,08 \times 10^{-14}$	<i>FGF5</i>
	rs12656497	5:32831939	T/C	0,43	0,90 (0,87–0,93)	$8,09 \times 10^{-11}$	<i>NPR3; AC026703.1</i>
Rasedus- hüpertensioon	rs7139122	12:53566790	G/A	0,98	0,71 (0,64–0,79)	$1,83 \times 10^{-10}$	<i>CSAD</i>
	rs167479	19:11526765	G/T	0,55	1,10 (1,06–1,13)	$8,85 \times 10^{-9}$	<i>RGL3</i>
	rs2208589	20:47408414	A/G	0,14	0,85 (0,81–0,89)	$1,29 \times 10^{-10}$	<i>PREX1</i>
	rs74715624	20:57753203	T/C	0,83	0,86 (0,82–0,89)	$1,91 \times 10^{-13}$	<i>ZNF831</i>

2.3.2. Kandidaatgeenide prioritseerimine FUMA-s

Riskilookuste kõige tõenäolisemate kandidaatgeenide kaardistamiseks kasutati FUMA pakutud võimalusi ning andmekihte. Arvesse võeti järgnevaid tegureid:

1. juhtivale SNP-le lähimad geenid;
2. MAGMA geeni-põhise analüüsi tulemused;
3. riskilookuses olevate variantide mõju geeniekspressioonile (nn eQTL);
4. riskilookuses olevate variantide ülekatet teadaolevate kromatiininteraktsioonidega.

eQTL ja kromatiininteraktsioonidel põhinevate geenikaardistamiste tulemused on kantud FUMA platvormi abil *circos*'e joonistele (vt Joonis 6 ja Lisa 2). Ülevaade kandidaatgeenide kaardistamisest on toodud välja Tabelis 5. Potentsiaalsete PE ja GH tekkemehhanismidesse kaasatud geenide identifitseerimiseks keskenduti nendele kandidaatgeenidele, mida kaardistati vähemalt 2 andmekihi abil. Samuti pöörati tähelepanu geenidele, milles paiknesid mittesünonüümsed kandidaatvariandid. Lõpuks võrreldi prioritseeritud geenide nimekirja varasemate preeklampsia ja rasedushüpertensiooni seoseuringutega.

Preeklampsia GWAS metaanalüüsi tulemuste põhjal kaardistas MAGMA geenipõhine analüüs 4 geeni. *PSORSIC2* geen 6. kromosoomis ning *ATXN2*, *BRAP* ja *ACAD10* geenid 12. kromosoomis. Kuna nendes kromosoomides üksikvariantide analüüsis riskilookuseid ei tuvastatud, siis ei uuritud sealsete geneetiliste variantide eQTL ja kromatiininteraktsioone. Preeklampsia ainsa riskilookusega (4. kromosoom, juhtiv SNP rs16998073) seostati 14 unikaalset geeni, millest 4 tulid välja eQTL ja kromatiininteraktsioonide põhisel kaardistamisel: *BMP3*, *C4orf22*, *FGF5*, *PRDM8*. *FGF5* on ühtlasi ka GWAS signaalile lähim geen.

Rasedushüpertensiooni esimese riskilookuse (1. kromosoom, juhtiv SNP rs143439093) abil prioritseeriti 19 geeni, neist 7 geeni nii eQTL kui kromatiininteraktsiooni abil: *MIIP*, *MFN2*, *PLOD1*, *CLCN6*, *MTHFR*, *AGTRAP* ja *FBXO6*. *KIAA2013* geeni kohta puudus kromatiininteraktsiooni andmekiht, kuid see oli juhtiva SNP-i lähim geen. Eelmises alapeatükis (2.3.1. Geneetiliste variantide prioritseerimine FUMA-s) kirjeldati riskilookuse sõltumatu SNP-i potentsiaalset negatiivset mõju *CLCN6* geenile.

Teise riskilookuse (4. kromosoom, juhtiv SNP rs16998073) kaardistatud geenid katusid preeklampsia riskilookuse geenidega täies ulatuses.

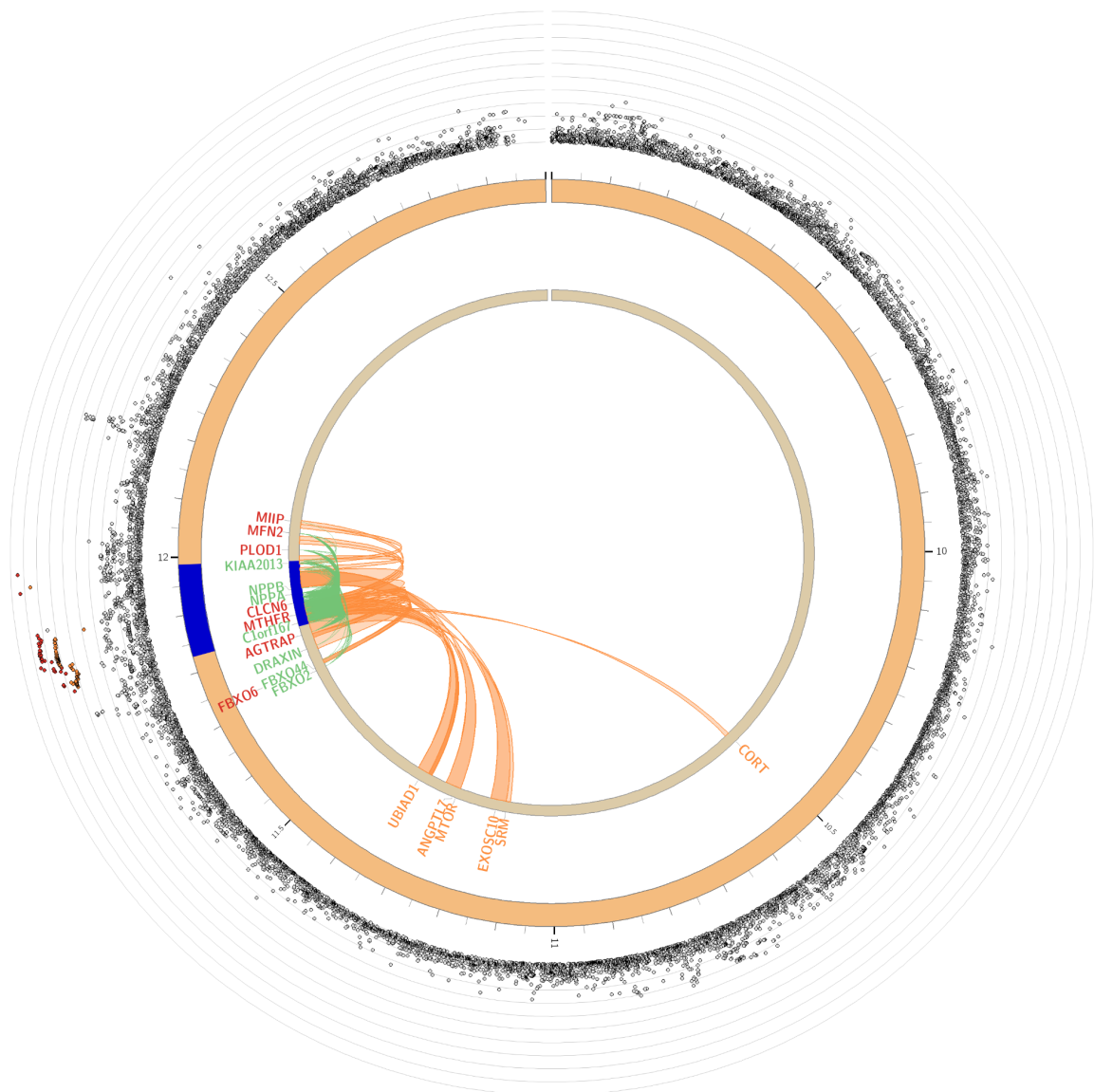
Kolmandas riskilookuses (5. kromosoom, juhtiv SNP rs12656497) õnnestus kaardistada 8 geeni. Tõendeid kromatiininteraktsioonide ja eQTL andmebaasidest oli kahel geenil (*AC026703.1* ja *NPR3*), mis olid samuti ka lookuse juhtiva SNP-i lähimad geenid.

Neljandas lookuses (12. kromosoom, juhtiv SNP rs7139122) oli kõige rohkem kaardistatuid geene – 49. Nendest 6 kaardistati eQTL ja kromatiininteraktsioonide abil: *MFSD5*, *ITGB7*, *IGFBP6*, *SPRYD3*, *TENCI* ja *KRT72*. Lisaks, juhtiva SNP-i lähim geen *CSAD* kaardistati kromatiininteraktsioonide abil.

Viies riskilookus (19. kromosoom, juhtiv SNP rs167479) oli seotud 5 geeniga, kuid ükski geen ei olnud kaardistatud nii eQTL kui kromatiininteraktsioonide informatsiooni põhjal. Sellegipoolest tuvastati eQTL informatsiooni *RGL3* geeni jaoks. Kuna lisaks sellele oli teada riskilookuse juhtiva SNP-i negatiivset mõju *RGL3* geenile, on see meie hinnangul selle piirkonna kõige tõenäolisem kandidaatgeen.

Kuuendas riskilookuses (20. kromosoom, juhtiv SNP rs2208589) on üks kandidaatgeen – *PREX1*, mis on signaalile lähim geen, kuid lisaks kaardistati ka eQTL info põhjal. Seitsmendas riskilookuses (20. kromosoom, juhtiv SNP rs74715624) oli pakutud 20 kandidaatgeeni. Kromatiininteraktsioonide ja eQTL andmete alusel jäi sõelale 4 kandidaatgeeni: *EDN3*, *ZNF831*, *SLMO2* ja *CTSZ*.

Kokku oli tuvastatud 27 potentsiaalset kandidaatgeeni, millest kuue geeni (*CLCN6*, *MTHFR*, *FGF5*, *CSAD*, *PREX1*, *ZNF831*) võimalikud seosed on kirjeldatud varasemates preeklampsia ja rasedushüpertensiooni GWAS-ides (vt Tabel 1, peatükk 1.4. Preeklampsia ja rasedushüpertensiooni senised ülegenoomsed seoseuuringud).



Joonis 6. Rasedushüpertensiooni GWAS metaanalüüsi tulemuste põhjal valmistatud 1. kromosoomi *circos*'e joonis. Välimisele kihile on kantud Manhattani joonis. Joonisele on kantud geneetilised variandid, mille p -väärtus $\leq 0,05$. Y-teljel on vastava geneetilise variandi p -väärtuse negatiivne kümnendlogaritm. SNP-d on tähistatud vastavalt nende suurimale r^2 väärtusele ühega riskilookuses olevatest sõltumatutest olulistest SNP-dest: punane ($r^2 > 0,8$), oranž ($r^2 > 0,6$), roheline ($r^2 > 0,4$), sinine ($r^2 > 0,2$) ja hall ($r^2 \leq 0,2$). Keskmisele kihile on kantud kromosoom koos koordinaatidega, sisemisele kihile – koordinaatideta; sinise värviga on tähistatud riskilookus. Kandidaatgeenid on tähistatud vastavalt toetusandmete olemusele: roheline (eQTL), oranž (kromatiininteraktsioon) ja punane (eQTL ja kromatiininteraktsioon). Sisemise kihi kaared tähistavad eQTL-e (roheline) ja kromatiininteraktsioone (oranž). Joonis on tehtud FUMA veebiplatvormi abil.

Tabel 5. Kandidaatgeenide tabel. Paksus kirjas on välja toodud geenid, mida kaardistati vähemalt kahe andmekihi abil.

Diagnoos	Marker	Lähim geen	MAGMA geenipõhine analüüs	eQTL	Kromatiininteraktsioon
Preeklampsia	rs16998073	<i>FGF5</i>	–	<i>BMP3, C4orf22, FGF5, PRDM8</i>	<i>THAP9, SEC31A, HNRNPD, PRKG2, BMP3, C4orf22, FGF5, PRDM8, ANTXR2, GK2, NAA11, PAQR3, ANXA3, FRAS1</i>
Rasedushüper-tensioon	rs143439093	<i>KIAA2013</i>	<i>CLCN6, MTHFR</i>	<i>MIIP, MFN2, PLOD1, KIAA2013, NPPB, NPPA, CLCN6, MTHFR, C1orf167, AGTRAP, DRAXIN, FBXO44, FBXO6, FBXO2</i>	<i>MIIP, MFN2, PLOD1, CLCN6, MTHFR, AGTRAP, FBXO6, UBA1AD1, ANGPTL7, MTOR, EXOSC10, SRM</i>
	rs16998073	<i>FGF5</i>	–	<i>BMP3, C4orf22, FGF5, PRDM8</i>	<i>THAP9, SEC31A, HNRNPD, PRKG2, BMP3, C4orf22, FGF5, PRDM8, ANTXR2, GK2, NAA11, PAQR3, ANXA3, FRAS1</i>
	rs12656497	<i>NPR3, AC026703.1</i>	<i>AC026703.1</i>	<i>AC026703.1, NPR3</i>	<i>AMACR, SLC45A2, RXFP3, ADAMTS12, TARS, AC026703.1, NPR3, ZFR</i>
	rs7139122	<i>CSAD</i>	–	<i>MFSD5, ITGB7, IGFBP6, SPRYD3, TENC1, KRT1, KRT73, KRT72</i>	<i>ESYT1, PA2G4, RPL41, ZC3H10, RP11-603J24.9, APT5G2, ATF7, RP11-793H13.10, TARBP2, MAP3K12, NPFF, PRR13, PCBP2, AMHR2, SP1, C12orf10, SP7, ESPL1, PFDN5, AAAS, MFSD5, RARG, ZNF740, CSAD, ITGB7, IGFBP6, SOAT2, SPRYD3, TENC1, EIF4B, KRT78, KRT72, KRT74, KRT5, KRT6A, KRT6B, KRT6C, KRT75, KRT84, KRT82, KRT85, AC121757.1, KRT81, KRT83, KRT86, KRT7, C12orf44</i>
	rs167479	<i>RGL3</i>	–	<i>ZNF627, ACP5, RGL3, AC011475.1</i>	<i>SWSAP1</i>

rs2208589	<i>PREX1</i>	<i>PREX1</i>	<i>PREX1</i>	–
rs74715624	<i>ZNF831</i>	<i>ZNF831</i>	<i>EDN3, ZNF831, SLMO2, CTSZ, NELFCD</i>	<i>EDN3, ZNF831, TUBB1, SLMO2, CTSZ, ATP5E, GNAS, STX16-NPEPL1, STX16, MGC4294, VAPB, ANKRD60, C20ORF85, ZBP1, PCK11, CTCFL, MTRNR2L3, RAE1, SPOII</i>

2.3.3. LD skoori regressioon

Diagnooside jagatud geneetika paremaks mõistmiseks, arvutati LD skoori regressiooni abil hinnangulised päritavuskoeffitsiendid ($h_{SNP L}^2$) ja geneetilise korrelatsiooni kordajad (r_g). Sünnitusabi ning günekoloogilistele diagnoosidele keskendunud analüüsi kaasati 47 raseduse ja naiste tervisega seotud diagnoosi. Kattuva fenotüübiga diagnoosid ja diagnoosid, mille geneetiliste korrelatsioonide hindamine ebaõnnestus valimi väikese suuruse (juhtude arv jääb vahemikku 458–1585, mediaan 1054) ning madala päritavushinnangu ($h_{SNP L}^2$ jääb vahemikku -0,0525–0,0376, mediaan -0,02435) tõttu, eemaldati analüüsist. Filtreerimise järgselt jäi uuringusse 38 tunnust.

2.3.3.1. Päritavuse hinnang

LD skoori regressiooni abil arvutatud (h_O^2) ja riskiskaalale teisaldatud (h_L^2) päritavuskoeffitsiendid on välja toodud Tabelis 6. Nii preeklampsia kui rasedushüpertensiooni riskiskaala pärilikkus jääb 10% kanti. Uuritud diagnoosid omavad ühtesid suuremaid geneetilisi komponente rasedusaegsete ning muude naiste tervisega seotud diagnooside hulgas (vt Lisa 5).

Tabel 6. Preeklampsia ja rasedushüpertensiooni päriliku komponendi väärtused ja standardvead.

Diagnoos	$h_{SNP O}^2$ (SE)	$h_{SNP L}^2$ (SE)
Preeklampsia	0,015 (0,0023)	0,1146 (0,0176)
Rasedushüpertensioon	0,0235 (0,0023)	0,1449 (0,0142)

2.3.3.2. Geneetilised korrelatsioonid

Järgnevalt vaadeldi 38 reproduktiivtunnuse vahelisi geneetilisi seoseid (vt Lisa 3).

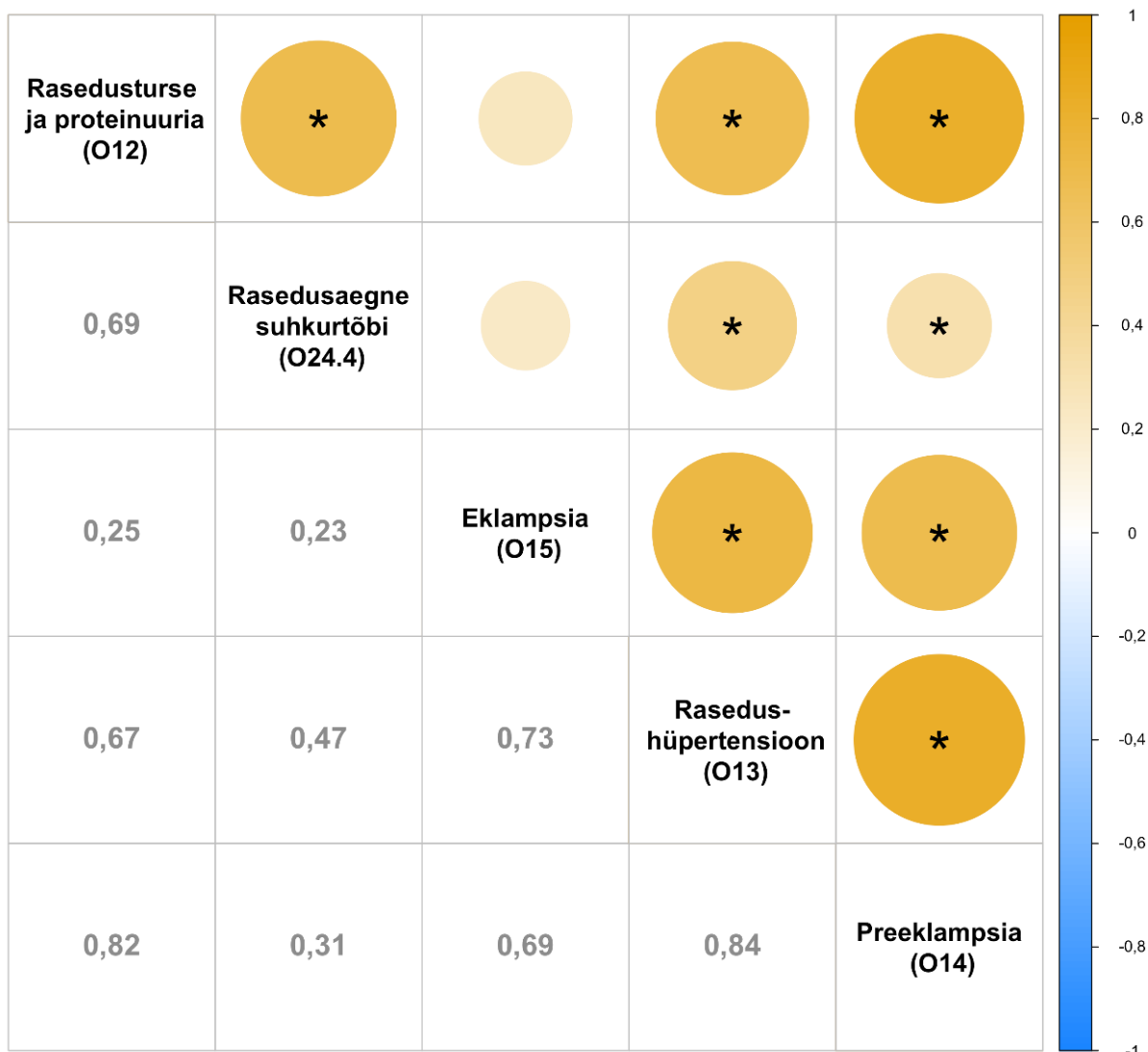
Pärast diagnooside sarnasuse alusel järjestamist ilmnes selge klaster, kuhu lisaks preeklampsia ja rasedushüpertensioonile kuulusid ka eklampsia, rasedusturse ja proteiinuuria ning raseduse ajal tekkiv suhkurtõbi. Nende tunnuste vahelised seosed on eraldi visualiseeritud (vt Joonisel 7).

Ootuspäraselt on leitud tugev positiivne seos preeklampsia ja rasedushüpertensiooni vahel ($r_g = 0,84$, SE = 0,07, q-väärtus = $2,34 \times 10^{-35}$) ning preeklampsia ja selle tüsistuse, eklampsia, vahel ($r_g = 0,69$, SE = 0,26, q-väärtus = $3,07 \times 10^{-2}$). Samuti täheldati tugevat positiivset geneetilist korrelatsiooni PE ning hüpertensioonita rasedusturse ja proteiinuuria vahel ($r_g = 0,82$, SE = 0,14, q-väärtus = $7,43 \times 10^{-8}$). Preeklampsia ja rasedusaegse suhkurtõve vahel ilmnes

mõõdukas positiivne seos ($r_g = 0,31$, $SE = 0,09$, q -väärtus = $3,17 \times 10^{-3}$). Suuremast analüüsist, mis hõlmas 38 reproduktiivtunnust, ei ilmunud ühtegi statistiliselt olulist negatiivset seost PE-ga.

Rasedushüpertensiooni positiivsed seosed on täheldatud järgmiste klastrisiseste tunnustega: eklampsia ($r_g = 0,73$, $SE = 0,26$, q -väärtus = $2,14 \times 10^{-2}$), hüpertensioonita rasedusturse ja proteiinuuria ($r_g = 0,67$, $SE = 0,11$, q -väärtus = $6,87 \times 10^{-8}$), rasedusaegne suhkurtõbi ($r_g = 0,47$, $SE = 0,06$, q -väärtus = $3,41 \times 10^{-11}$) ning eelpool mainitud seos preeklampsiaga. Statistiliselt olulised negatiivsed seosed GH ja uuringusse kaasatud reproduktiivtunnuste vahel puuduvad.

Lisaks tuvastati 3 diagnoosi, mis on statistiliselt oluliselt seotud nii preeklampsia kui rasedushüpertensiooniga (r_{gPE} ja r_{gGH}): enneaegne sünnitus ($r_{gPE} = 0,3$, $r_{gGH} = 0,3$), iseeneslik abort ($r_{gPE} = 0,54$, $r_{gGH} = 0,49$) ning munajuha- ja munasarjapõletik ($r_{gPE} = 0,35$, $r_{gGH} = 0,32$). Statistiliselt oluline seos preeklampsiaga, kuid mitte rasedushüpertensiooniga, oli emakakaela põletikulisel haigusel ($r_g = 0,45$, q -väärtus = $4,57 \times 10^{-2}$) ja korratul menstruatsioonil ($r_g = 0,28$, q -väärtus = $1,14 \times 10^{-2}$). Analüüsi käigus tuvastati ka diagnoosid, mis olid statistiliselt oluliselt seotud rasedushüpertensiooniga, kuid mitte preeklampsiaga: rohke, sage ja korratu menstruatsioon ($r_g = 0,18$, q -väärtus = $2,11 \times 10^{-2}$) ning varane rasedusaegne verejooks ($r_g = 0,26$, q -väärtus = $3,79 \times 10^{-2}$).



Joonis 7. Geneetilised korrelatsioonid 5 raseduskomplikatsiooni vahel. Kollase värviga on tähistatud positiivsed seosed, sinise värviga on tähistatud negatiivsed seosed (puuduvad). Diagonaalis on toodud välja diagnooside nimetused ja vastavad RHK-10 tähtnumbrilised koodid. Ülemises kolmnurgas on visualiseeritud tunnuste vahelised geneetilised korrelatsioonid. Ringide suurus peegeldab r_g absoluutväärtust. Tärniga („*“) on tähistatud statistiliselt olulised (korrigeeritud p-väärtus < 0.05) seosed. Alumises kolmnurgas on esitatud seoste numbrilised väärtused.

Kasutades CTG-VL veebiplatvormi uuriti geneetilisi seoseid preeklampsia ja rasedusaegse hüpertensiooni ning 1403 veebiplatvormil kättesaadava tunnuse vahel.

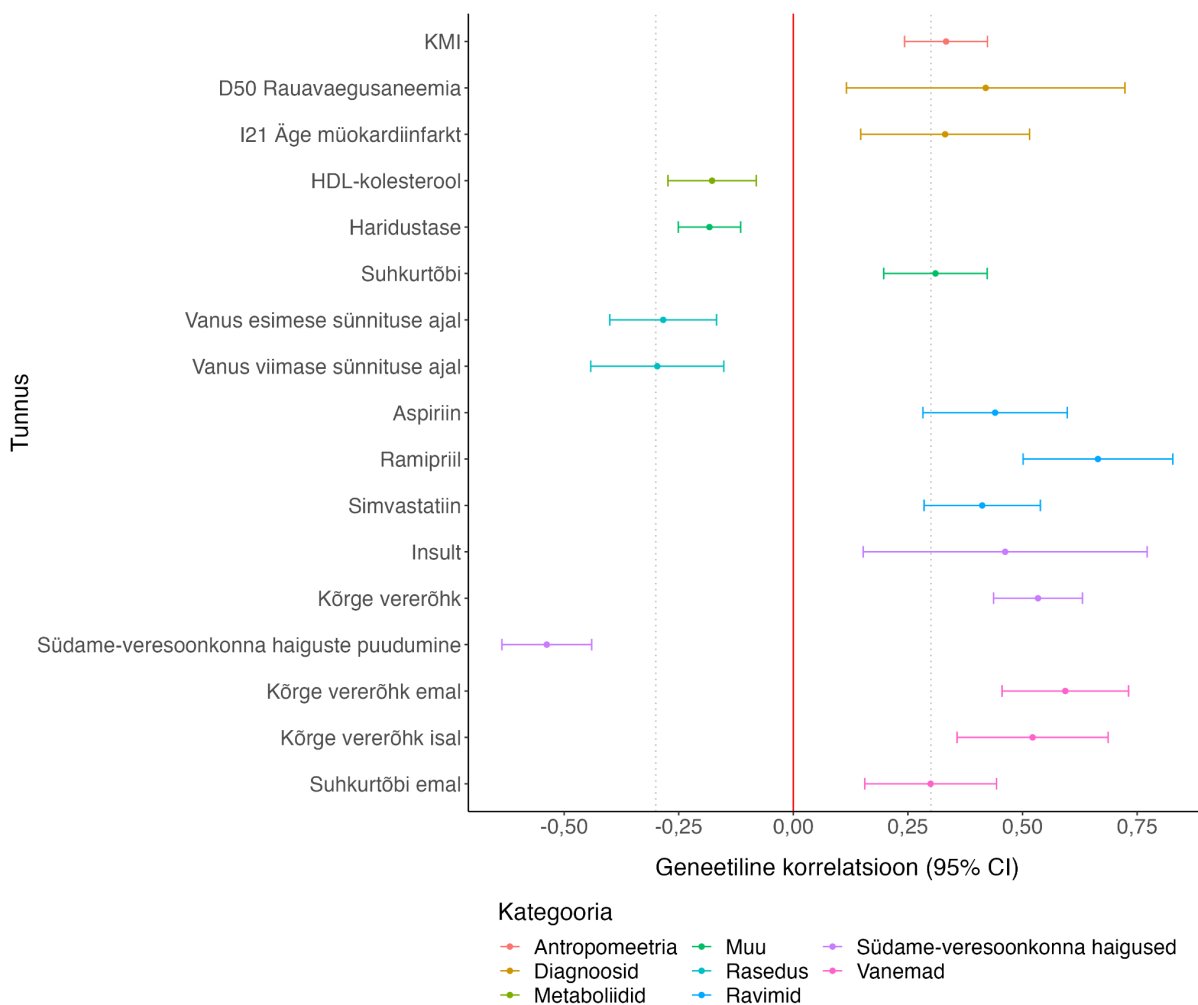
Preeklampsia geneetilisi korrelatsioone oli võimalik edukalt hinnata 1359 tunnuse puhul. Pärast p-väärtuste korrigeerimise, ületas 189 korrelatsiooni statistilise olulisuse piirnivoo (korrigeeritud p-väärtus ehk q-väärtus < 0,05). 74 korrelatsiooni jäid -0,3 kuni 0,3 vahemikust välja, neid peeti mõõdukateks. Nendest vaid 6 geneetilise korrelatsiooni hinnangulist väärtust olid 0,6-st suurem või -0,6-st väiksem, nendest 5 on korrelatsioonid ravimitega (furosemiid, ramipriil, enalapriil, perindopriil, lisinopriil) ja infarkt sugulastel ($r_g = 0,43$, q-väärtus = $1,93 \times 10^{-2}$). Huvitavamate

tulemuste kokkuvõte on esitatud joonisel (vt Joonis 8). Ootusepäraselt täheldati positiivseid geneetilisi seoseid preeklampsia ja südame-veresoonkonna haiguste vahel ning eelmainitud haiguste puudumine on preeklampsiaga mõõdukas negatiivses korrelatsioonis. HDL-kolesteroolil on samuti negatiivne seos ($r_g = -0.18$, q -väärtus = 4×10^{-3}) preeklampsiaga, kuid hinnanguline väärtus on väga nõrk. Huvitavalt on antud analüüsis kirjeldatud preeklampsia geneetiline korrelatsioon kehamassiindeksiga madal ($r_g = 0,33$, q -väärtus = $6,17 \times 10^{-11}$).

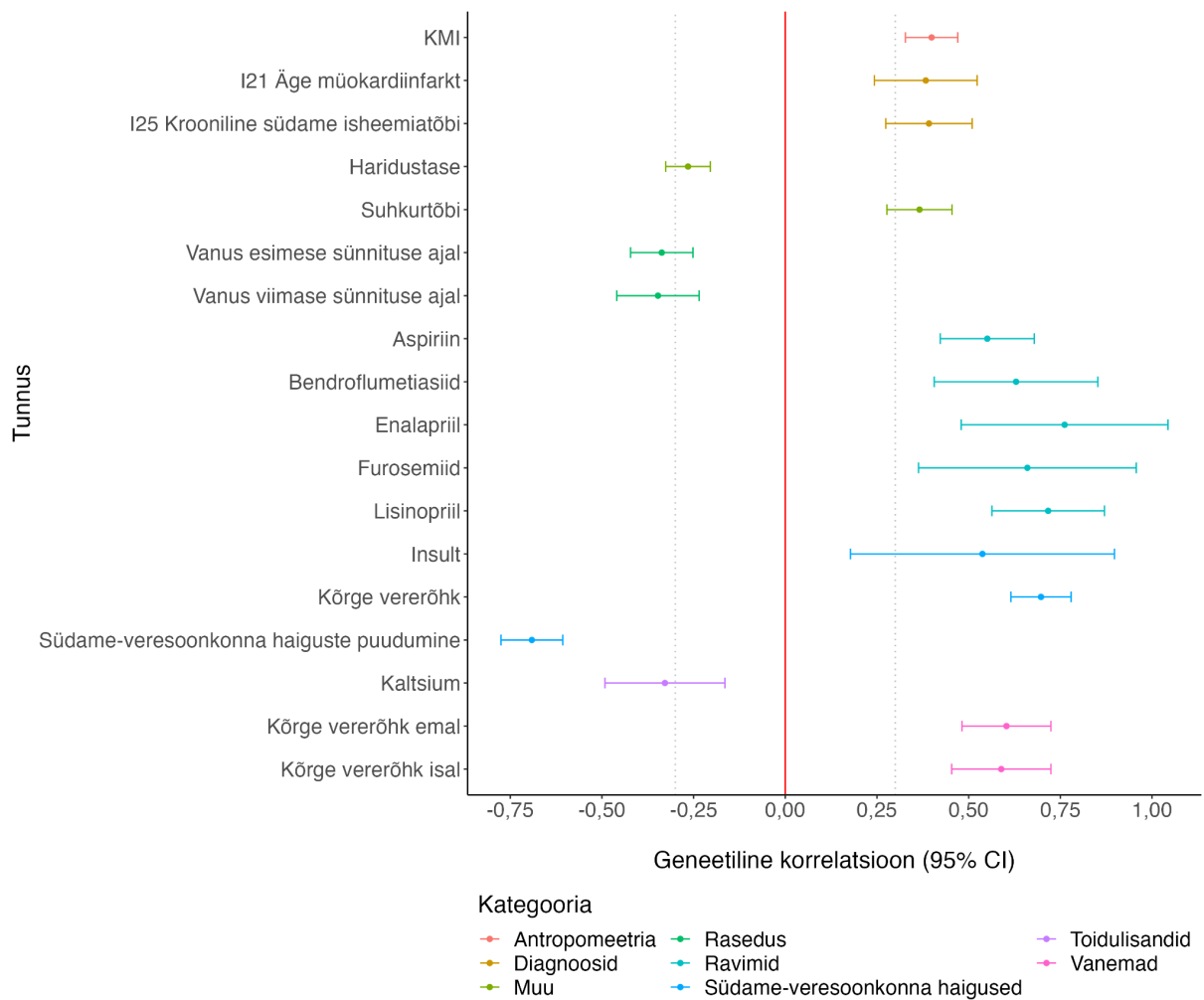
Sarnaselt PE-le sai antud analüüsi käigus tuvastada geneetilisi korrelatsiooni gestatsiooni hüpertensiooni ja 1358 tunnuse vahel. Tulemuste korrigeerimise tagajärjel ületasid statistilise olulisuse nivoo 328 tunnust. Nõrga korrelatsiooni väärtuste vahemikust (-0,3...0,3) jäi välja 137 tunnust ning 19 seost peeti tugevateks. Tugevad seosed ilmneseid GH ja vererõhu alandavate ravimite ning kõrgvererõhutõve vahel. Huvitavamad tulemused on näidatud joonisel (vt Joonis 9).

Preeklampsia ja rasedushüpertensiooni tulemuste võrdlemisel selgus, et 10 PE-ga mõõdukas korrelatsioonis olnud tunnust (r_g absoluutväärtus $\geq 0,3$) ei olnud mõõdukas seoses rasedushüpertensiooniga. Tunnuseid vaadati manuaalselt üle ning sarnaseid tunnuseid (nt rauavaegusaneemia ja D50 rauavaegusaneemia) grupeeriti. Sotsiaalmajanduslike tunnuseid kõrvale jättes, on vaid PE-ga mõõdukalt seotud järgmised tunnused: katarakt ($r_g = 0,3668$, q -väärtus = $1,70 \times 10^{-2}$), sugulaste mittejuhuslik surm ($r_g = 0,4308$, q -väärtus = $1,93 \times 10^{-2}$), spinaalstenoos ($r_g = 0,332$, q -väärtus = $4,27 \times 10^{-2}$), D50 rauavaegusaneemia ($r_g = 0,4196$, q -väärtus = $4,91 \times 10^{-2}$) ja N39 kuseelundite muud haigusseisundid ($r_g = 0,3375$, q -väärtus = $2,82 \times 10^{-2}$).

Rasedushüpertensiooniga, kuid mitte preeklampsiaga, mõõdukas korrelatsioonis oli 73 tunnust. Lähemal vaatlusel selgus, et paljud korrelatsioonid olid seotud südameveresoonkonna haigustega (näiteks hüpertensioon, kardiovaskulaarsed haigused sugulastel) ning vererõhku ja vere kolesterooli alandavate ravimitega (näiteks valsartan, esetimiib). Samuti ilmnese mitmel erineval kujul seos suhkurtõvega (suhkurtõbi emal, teist tüüpi suhkurtõbi).



Joonis 8. Geneetilised korrelatsioonid preeklampsia ja teiste tunnuste vahel Euroopa päritoluga isikutel. Hinnanguline geneetilise korrelatsiooni väärtus (r_g) on kantud joonisele punktina. Veajoon tähistab 95%-usaldusintervalli (95% CI). Hall katkendjoon eristab geneetilisi korrelatsioone, mis jäävad vahemikku $-0,3 \dots 0,3$ ning on peetud nõrkadeks. Punane pidevjoon näitab korrelatsiooni puudumist. KMI – kehamassiindeks.



Joonis 9. Geneetilised korrelatsioonid rasedusaegse hüpertensiooni ja teiste tunnuste vahel Euroopa päritoluga isikutel. Hinnanguline geneetilise korrelatsiooni väärtus (r_g) on kantud joonisele punktina. Veajoon tähistab 95%-usaldusintervalli (95% CI). Hall katkendjoon eristab geneetilisi korrelatsioone, mis jäävad vahemikku -0,3...0,3 ning on peetud nõrkadeks. Punane pidevjoon näitab korrelatsiooni puudumist. KMI – kehamassiindeks.

2.4. Arutelu

Antud magistritöö üheks eesmärgiks oli kirjeldada preeklampsia ja rasedushüpertensiooni geneetilisi riskifaktoreid. Selleks kaardistati GWAS metaanalüüsi tulemusi, kirjeldati diagnoosidega seotuid geneetilisi riskilookuseid ning pakuti välja kandidaatgeene. Metaanalüüsi kasuks otsustati lähtudes tõsiasjast, et uuritavate rasedustüsistustega diagnoositud isikute arv TÜ Eesti Geenivaramus on liiga madal ning ei võimalda tagada piisavat uuringu võimsust. Lisaks on varasemad uuringud näidanud, et Soome ja Eesti andmete kombineerimine võib aidata tuvastada seotud variante, mida teistes populatsioonides esineb harvem (Tyrmi *et al.*, 2022). Uuringu käigus leiti, et preeklampsiaga on assotsieerunud üks lookus ning rasedushüpertensiooniga on seotud seitse lookust, millest üks kattub preeklampsia uuringus ainsa tuvastatud signaaliga. Uurimistöö teises osas hinnati LD skoori regressiooni meetodi abil preeklampsia ja rasedushüpertensiooni päritavuskoeffitsiendid, mis mõlema tunnuse puhul jäid 10–15% vahele. Järgnevalt hinnati uuritavate diagnooside jagatud geneetilist tausta omavahelises võrdluses ning 36 sünnitusabi ning günekoloogilise diagnoosiga. Leitud geneetilised korrelatsioonid viitavad sellele, et preeklampsia ja rasedushüpertensiooni vahel on jagatud geneetikat. Veebiplatvorm CTG-VL võimaldas teostada sarnast analüüsi veel mitmete teiste fenotüüpide ja tunnustega (diagnoosid, retseptiravimid, antropomeetrilised näitajad, sotsiaalmajanduslik staatus).

Ülegenoomse assotsiatsiooniuuringu abil leiti ühe olulise seose preeklampsia ja 4. kromosoomis paikneva geneetilise variandi rs16998073 vahel, mis on samaaegselt ka rasedushüpertensiooniga seotud variant. rs16998073 paikneb *FGF5* geeni lähedal ning antud geeni seos kõrge vererõhuga on pikemat aega teada (Liu *et al.*, 2011). Hiljutine Xin jt uuring kinnitas, et rs16998073 variandi T-alleeli esinemissagedus Aasia päritoluga PE-ga diagnoositud naiste seas on kõrgem. Uuringus näidati katseliselt, et T-alleel soodustab *FGF5* ekspressiooni. Lisaks selgus, et *FGF5* ekspressioonitase PE-ga diagnoositud naiste platsentas on kõrgem (Xin *et al.*, 2022).

20. kromosoomis leiti ka teine seotud lookus (rs2208589), mis paikneb *PREX1* geeni intronis. Antud variandi ja rasedushüpertensiooni vaheline seos on leitud ka FinnGen R6 andmete põhjal teostatud ülegenoomses assotsiatsiooniuuringus, kuid FinnGen ja Ühendkuningriigi biopanga metaanalüüsis seost replitseerida ei õnnestunud (p -väärtus = 2.7×10^{-4}) (Changalidis *et al.*, 2022). Ei ole välistatud, et tegu võib olla Soome populatsiooni-spetsiifilise leiuga, kuna erinevate Euroopa populatsioonide lõikes võib täheldada mõningasi alleelisageduse erinevusi (harvema alleeli sagedus Soomes 0.11, Eestis 0.16, Euroopas keskmiselt 0.20). Varasemalt on

spekuleeritud, et *PREX1* poolt kodeeritud valk soodustab rinnavähi teket ning on näidatud, et *PREX1* võib kaudselt mõjutada endoteeli barjäärifunktsiooni (Montero *et al.*, 2011; Reinhard *et al.*, 2021).

Üheks huvitavamaks leiuks võib pidada geneetilist varianti rs13306561, mis asub *MTHFR* geeni transkripti 5'-mittetransleeritavas piirkonnas. *MTHFR* geeni polümorfismid on teadaolevalt seotud valgu aktiivsusega, mida saab kaudselt väljendada homotsüsteiini taseme kaudu. On näidatud, et homotsüsteiini taseme tõus on omakorda seotud vererõhu tõusuga (Kumar *et al.*, 2005; Weng *et al.*, 2020).

Uuringu teises osas teostati preeklampsia ja rasedushüpertensiooni LD skoori regressiooni tulemuste analüüsi. Meetodi abil hinnati tunnuste päritavust, ning leiti, et preeklampsia päritavus on 11.1% ja rasedushüpertensiooni päritavus 14.5%. Kui rasedushüpertensiooni päritavuse hinnang on varasemalt kirjeldatud päritavusega – 20% – võrreldav, siis preeklampsia päritavuskoeffitsient on varasemalt kirjeldatud 38%-st palju madalam (Nilsson *et al.*, 2004; Steinhorsdottir *et al.*, 2020). Erinevus varasemalt kirjeldatud ning antud magistritöös saadud SNP-põhisest päritavusest vajab edasist uurimist. Huvitav on asjaolu, et päritavuskoeffitsiendid olid saadud kasutades sama meetodit (LD skoori regressioon), seega on erinevused tõenäoliselt tingitud erinevast fenotüübidefinitsioonist, heterogeensemast valimist, või väiksemast valimisuurusest.

Lisaks vaadeldi uuringu teise osa raames geneetilisi korrelatsioone preeklampsia ja rasedushüpertensiooni ning 36 sünnitusabi ja günekoloogia diagnoosiga. Uuringust ilmselge assotsiatsioon preeklampsia, rasedushüpertensiooni, eklampsia, hüpertensioonita rasedusturse ja proteiinuuria ning rasedusdiabeediga. Nii preeklampsia kui rasedushüpertensioon näitasid lisaks positiivset seost kolme diagnoosiga: enneaegne sünnitus, iseeneslik abort ning munajuha- ja munasarjapõletik. Unikaalseid seoseid, mis olid omased vaid preeklampsiale või rasedushüpertensioonile oli 4, kuid nendes vaid emakakaela põletikulise haiguse korrelatsioon PE-ga on mõõduka suurusega. CTG-VL veebiplatvormi abil arvatati geneetilised korrelatsioonid üle ligikaudu 1400 fenotüübiga, mis aitasid veelgi enam kinnitada PE ja GH jagatud geneetikat, sest tulemused olid suuresti sarnased. Eraldi võiks siinkohal välja tuua mõõduka positiivse geneetilise korrelatsiooni kehamassiindeksiga, mis võib viidata sellele, et kehakaalu alandamine võib aidata vähendada preeklampsia ja rasedushüpertensiooni riski. Edasiste uuringute käigus tuleks lisaks selgitada, kuivõrd mõjutab saadud tulemusi see, et

preeklampsia ja rasedushüpertensiooni sümptomid on kohati kattuvad ning ei saa välistada valediagnoose, mis muudab uuritud fenotüübid mõnevõrra heterogeenseks.

Läbiviidud uuringu kitsaskohaks on asjaolu, et kasutatud valimite suurus on väike ning ei võimalda tuvastada kõigi preeklampsia ja hüpertensiooni tekkemehhanismidega assotsieerunud üksiknukleotiidseid variante. Tulevikus võiks metaanalüüsi võimsuse suurendamiseks kaasata teisi kohorte, võimalusel ka teistest populatsioonidest, mis võimaldaks kirjeldada geneetilisi riskifaktoreid ka teistes populatsioonides ning seega parandada meie arusaama haiguse tekkemehhanismidest. Teiseks oluliseks aspektiks on see, et uuringu käigus vaadeldi vaid emapoolseid geneetilisi variante. Rasedust mõjutavad nii ema, isa kui ka lapse genoomid ning nende mõju ei ole ühtlane. On näidatud, et ühe ja sama geneetilise variandi seos preeklampsiaga varieerub vastavalt sellele, kes on efektilleeli kandjaks – nii on rs4769613 C-alleeli esinemine lapsel (platsenta), kuid mitte emal, seotud preeklampsia tekkega (Kikas *et al.*, 2020). Edaspidises töös saaks teostada sarnase ülesehitusega analüüsi preeklampsiaga ja/või rasedushüpertensiooni diagnoositud naiste laste geneetilise informatsiooni põhjal.

Lähtudes teaduskirjandusest on preeklampsia heterogeenne rasedustüsistus. Kuigi preeklampsiat jagatakse varajaseks ja hiliseks vormiks ning oletatakse, et kahe vormi tekkemehhanismid on erinevad, koondatakse preeklampsia erinevaid vorme RHK-10 süsteemis ühe koodi alla. See on ka üks peamisi suurte populatsioonipõhiste biopankade kitsaskohti – puuduvad täpsemad kliinilised andmed fenotüübidefinitsiooni rafineerimiseks ning võimalikuks stratifitseerimiseks. Vastavate andmete lisandumisel (näiteks oleks võimalik andmeid linkida Meditsiinilise Sünniregistriga) saaks vaadata kas diagnooside stratifitseerimine trimestrite või raseduskestvuse järgi annaks assotsiatsiooniuuringus lisatulemusi.

Kokkuvõte

Käesoleva magistritöö eesmärgiks oli analüüsida rasedushüpertensiooni ja preeklampsia emapoolseid geneetilisi riskifaktoreid ning seoseid teiste fenotüüpidega, et paremini mõista uuritavate haiguste patogeneesi. Selleks kasutati Eesti Geenivaramu ning Soome FinnGen andmestikke, viidi läbi geneetiliste seoseuringute meta-analüüs ning annoteeriti saadud tulemusi FUMA veebiplatvormil. Lisaks analüüsiti rasedushüpertensiooni ja preeklampsia omavahelist jagatud geneetikat ning geneetilisi seoseid nii teiste naistehaiguste kui muude tunnustega.

Töö tulemusel leiti üks lookus, mis on seotud preeklampsiaga ning seitse lookust, mis olid seotud rasedushüpertensiooniga. Erinevate andmekihtide kombineerimisel kaardistati võimalikud kandidaatgeenid seotud lookustes ning kokku tuvastati 27 potentsiaalset kandidaatgeeni, millest kuue geeni (*CLCN6*, *MTHFR*, *FGF5*, *CSAD*, *PREX1*, *ZNF831*) võimalikud seosed on kirjeldatud varasemates preeklampsia ja rasedushüpertensiooni GWAS-ides.

Geneetiliste korrelatsioonide analüüsimisel leiti tugev positiivne seos preeklampsia ja rasedushüpertensiooni vahel ning preeklampsia ja selle tüsistuse, eklampsia, vahel. Samuti täheldati tugevat positiivset geneetilist korrelatsiooni PE ning hüpertensioonita rasedusturse ja proteiinuuria vahel. Nähtud seosed kinnitavad nende tunnuste sarnast geneetilist tausta, kuid edasised uuringud peaksid lisaks hindama seda, mil määral võis saadud leide mõjutada kohati sarnane sümptomaatika ning sellest tulenev ülekate erinevate diagnooside vahel. Geneetiliste korrelatsioonide analüüs ligikaudu 1400 fenotüübiga kinnitas taaskord PE ja GH jagatud geneetikat, sest tulemused olid suuresti sarnased.

The genetics of preeclampsia and gestational hypertension, and association with other diagnoses in the Estonian Biobank and FinnGen data

Valentina Rukins

Summary

Hypertensive disorders of pregnancy, including preeclampsia and gestational hypertension, are the second leading cause of death among pregnant women, accounting for 14% of maternal deaths worldwide (Say *et al.*, 2014). It is well-known that preeclampsia and gestational hypertension share common risk factors, including primiparity, multiple pregnancy, diabetes mellitus, and body mass index (Shen *et al.*, 2017; ülevaade Dimitriadis *et al.*, 2023). Several attempts have been made to describe preeclampsia and gestational hypertension's genetic risks, but exact biological mechanisms remain elusive (Steinthorsdottir *et al.*, 2020; Honigberg *et al.*, 2023; Tyrmi *et al.*, 2023).

The aim of this master's thesis was to analyze the maternal genetic risk factors of pregnancy hypertension and preeclampsia, and associations with other phenotypes, in order to better understand the pathogenesis of the studied phenotypes. For this purpose, data from the Estonian Biobank and the FinnGen project were used, a meta-analysis of genetic association studies was performed, and the obtained results were annotated on the FUMA web platform. In addition, the shared genetics between pregnancy hypertension and preeclampsia, and with other traits and diagnoses from publicly available databases, were analysed.

As a result of the conducted analyses, one locus was found to be associated with preeclampsia and seven loci were associated with pregnancy hypertension. By combining different data layers, potential candidate genes were mapped in the associated loci and a total of 27 potential candidate genes were identified, six of which (*CLCN6*, *MTHFR*, *FGF5*, *CSAD*, *PREX1*, *ZNF831*) have been described previously.

In the genetic correlations analysis, a strong positive association was found between preeclampsia and pregnancy hypertension, and between preeclampsia and its complication, eclampsia. A strong positive genetic correlation was also observed between PE and pregnancy edema and proteinuria without hypertension. The observed associations confirm the similar genetic background of these traits, but further studies should also assess the extent to which the

findings were influenced by sometimes similar symptomatology and the resulting overlap between different diagnoses. The analysis of genetic correlations with approximately 1400 phenotypes confirmed again the shared genetics of PE and GH, as the results were largely similar.

Tänuõnad

Eelkõige tahan tänada oma juhendajaid Triin Laiska ja Reedik Mäge kannatlikkuse, mõistvuse ja mitmekülgse abi eest. Samuti soovin tänada doktoranti Natália Pujol Gualdot teostatud metaanalüüside eest.

Täna ka oma ema, elukaaslast ja sõpru, kes vaatamata kõigile raskustele hoidsid minu meeleolu üleval.

Kasutatud kirjandus

Abalos, E., Cuesta, C., Carroli, G., Qureshi, Z., Widmer, M., Vogel, J., Souza, J., & on behalf of the WHO Multicountry Survey on Maternal and Newborn Health Research Network (2014). Pre-eclampsia, eclampsia and adverse maternal and perinatal outcomes: a secondary analysis of the World Health Organization Multicountry Survey on Maternal and Newborn Health. *BJOG-Int J Obstet Gy*, 121(s1): 14–24.

Abalos, E., Cuesta, C., Grosso, A. L., Chou, D. & Say, L. (2013). Global and regional estimates of preeclampsia and eclampsia: a systematic review. *Eur J Obstet Gynecol Reprod Biol*, 170(1): 1–7.

American College of Obstetricians and Gynecologists (2013). Hypertension in Pregnancy: Executive Summary. *Obstet Gynecol*, 122(5): 1122.

American College of Obstetricians and Gynecologists (2020). Gestational Hypertension and Preeclampsia: ACOG Practice Bulletin, Number 222. *Obstet Gynecol*, 135(6): e237–e260.

Andraweera, P. H. & Lassi, Z. S. (2019). Cardiovascular Risk Factors in Offspring of Preeclamptic Pregnancies—Systematic Review and Meta-Analysis. *The Journal of Pediatrics*, 208: 104-113.e6.

Auton, A., Abecasis, G. R., Altshuler, D. M., ... National Eye Institute, N. (2015). A global reference for human genetic variation. *Nature*, 526(7571): 68–74.

Browning, S. R. & Browning, B. L. (2007). Rapid and Accurate Haplotype Phasing and Missing-Data Inference for Whole-Genome Association Studies By Use of Localized Haplotype Clustering. *Am J Hum Genet*, 81(5): 1084–1097.

Bulik-Sullivan, B., Finucane, H. K., Anttila, V., ... Neale, B. M. (2015). An atlas of genetic correlations across human diseases and traits. *Nat Genet*, 47(11): 1236–1241.

Bulik-Sullivan, B. K., Loh, P.-R., Finucane, H., Ripke, S., Yang, J., Patterson, N., Daly, M. J., Price, A. L. & Neale, B. M. (2015). LD Score Regression Distinguishes Confounding from Polygenicity in Genome-Wide Association Studies. *Nat Genet*, 47(3): 291–295.

Changalidis, A. I., Maksiutenko, E. M., Barbitoff, Y. A., Tkachenko, A. A., Vashukova, E. S., Pachuliia, O. V., Nasykhova, Y. A. & Glotov, A. S. (2022). Aggregation of Genome-Wide

Association Data from FinnGen and UK Biobank Replicates Multiple Risk Loci for Pregnancy Complications. *Genes (Basel)*, 13(12): 2255.

Chen, W., Srinivasan, S. R., Yao, L., Li, S., Dasmahapatra, P., Fernandez, C., Xu, J. & Berenson, G. S. (2012). Low Birth Weight Is Associated With Higher Blood Pressure Variability From Childhood to Young Adulthood. *Am J Epidemiol*, 176(Suppl 7): S99–S105.

Dimitriadis, E., Rolnik, D. L., Zhou, W., ... Menkhorst, E. (2023). Pre-eclampsia. *Nat Rev Dis Primers*, 9(1): 1–22.

Espeche, W. G., Salazar, M. R., Minetto, J., ... Carbajal, H. A. (2022). Hypertension arising after 20 weeks of gestation: gestational hypertension or masked chronic hypertension? *J Hum Hypertens*, 1–5.

Gyselaers, W. (2022). Hemodynamic pathways of gestational hypertension and preeclampsia. *Am J Obstet Gynecol*, 226(2): S988–S1005.

Honigberg, M. C., Truong, B., Khan, R. R., ... Natarajan, P. (2023). Polygenic prediction of preeclampsia and gestational hypertension. *Nat Med*, 29(6): 1540–1549.

Huai, J., Lin, L., Juan, J., Chen, J., Li, B., Zhu, Y., Yu, M. & Yang, H. (2021). Preventive effect of aspirin on preeclampsia in high-risk pregnant women with stage 1 hypertension. *J Clin Hypertens (Greenwich)*, 23(5): 1060–1067.

Huang, C., Li, J., Qin, G., Liew, Z., Hu, J., László, K. D., Tao, F., Obel, C., Olsen, J. & Yu, Y. (2021). Maternal hypertensive disorder of pregnancy and offspring early-onset cardiovascular disease in childhood, adolescence, and young adulthood: A national population-based cohort study. *PLoS Med*, 18(9): e1003805.

Khera, A. V., Chaffin, M., Aragam, K. G., ... Kathiresan, S. (2018). Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet*, 50(9): 1219–1224.

Kifflen, M., Le, A., Mao, S., Lali, R., Narula, S., Xie, F. & Paré, G. (2022). Cost-Effectiveness of Polygenic Risk Scores to Guide Statin Therapy for Cardiovascular Disease Prevention. *Circulation: Genomic and Precision Medicine*, 15(5): e003423.

- Kikas, T., Inno, R., Ratnik, K., Rull, K. & Laan, M. (2020). C-allele of rs4769613 Near FLT1 Represents a High-Confidence Placental Risk Factor for Preeclampsia. *Hypertension*, 76(3): 884–891.
- Kircher, M., Witten, D. M., Jain, P., O’Roak, B. J., Cooper, G. M. & Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet*, 46(3): 310–315.
- Kumar, J., Das, S. K., Sharma, P., Karthikeyan, G., Ramakrishnan, L. & Sengupta, S. (2005). Homocysteine levels are associated with MTHFR A1298C polymorphism in Indian population. *J Hum Genet*, 50(12): 655–663.
- Kurki, M. I., Karjalainen, J., Palta, P., ... Palotie, A. (2023). FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature*, 613(7944): 508–518.
- Leeuw, C. A. de, Mooij, J. M., Heskes, T. & Posthuma, D. (2015). MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLoS Comput Biol*, 11(4): e1004219.
- Leitsalu, L., Haller, T., Esko, T., ... Metspalu, A. (2015). Cohort Profile: Estonian Biobank of the Estonian Genome Center, University of Tartu. *International Journal of Epidemiology*, 44(4): 1137–1147.
- Lisonkova, S. & Joseph, K. S. (2013). Incidence of preeclampsia: risk factors and outcomes associated with early- versus late-onset disease. *American Journal of Obstetrics and Gynecology*, 209(6): 544.e1-544.e12.
- Liu, C., Li, H., Qi, Q., Lu, L., Gan, W., Loos, R. J. & Lin, X. (2011). Common variants in or near FGF5, CYP17A1 and MTHFR genes are associated with blood pressure and hypertension in Chinese Hans. *J Hypertens*, 29(1): 70–75.
- Loh, P.-R., Danecek, P., Palamara, P. F., ... L Price, A. (2016). Reference-based phasing using the Haplotype Reference Consortium panel. *Nat Genet*, 48(11): 1443–1448.
- Mägi, R. & Morris, A. P. (2010). GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics*, 11(1): 288.
- Manolio, T. A. (2013). Bringing genome-wide association findings into clinical use. *Nat Rev Genet*, 14(8): 549–558.

- Mbatchou, J., Barnard, L., Backman, J., ... Marchini, J. (2021). Computationally efficient whole-genome regression for quantitative and binary traits. *Nat Genet*, 53(7): 1097–1103.
- Mitt, M., Kals, M., Pärn, K., ... Palta, P. (2017). Improved imputation accuracy of rare and low-frequency variants using population-specific high-coverage WGS-based imputation reference panel. *Eur J Hum Genet*, 25(7): 869–876.
- Mol, B. W. J., Roberts, C. T., Thangaratinam, S., Magee, L. A., de Groot, C. J. M. & Hofmeyr, G. J. (2016). Pre-eclampsia. *Lancet*, 387(10022): 999–1011.
- Montero, J. C., Seoane, S., Ocaña, A. & Pandiella, A. (2011). P-Rex1 participates in Neuregulin-ErbB signal transduction and its expression correlates with patient outcome in breast cancer. *Oncogene*, 30(9): 1059–1071.
- Nilsson, E., Salonen Ros, H., Cnattingius, S. & Lichtenstein, P. (2004). The importance of genetic and environmental effects for pre-eclampsia and gestational hypertension: a family study. *BJOG*, 111(3): 200–206.
- Pujol-Gualdo, N., Läll, K., Lepamets, M., Rossi, H.-R., Arffman, R. K., Piltonen, T. T., Mägi, R. & Laisk, T. (2022). Advancing our understanding of genetic risk factors and potential personalized strategies for pelvic organ prolapse. *Nat Commun*, 13: 3584.
- Reinhard, N. R., Van Der Niet, S., Chertkova, A., Postma, M., Hordijk, P. L., Gadella Jr., T. W. J. & Goedhart, J. (2021). Identification of guanine nucleotide exchange factors that increase Cdc42 activity in primary human endothelial cells. *Small GTPases*, 12(3): 226–240.
- Renlund, M. A.-K., Jääskeläinen, T. J., Kivelä, A. S. E., Heinonen, S. T., Laivuori, H. M. & Sarkola, T. A. (2023). Blood pressure, arterial stiffness, and cardiovascular risk profiles in 8-12-year-old children following preeclampsia (FINNCARE-study). *J Hypertens*, 41(9): 1429–1437.
- Roberge, S., Nicolaides, K. H., Demers, S., Villa, P. & Bujold, E. (2013). Prevention of perinatal death and adverse perinatal outcome using low-dose aspirin: a meta-analysis. *Ultrasound Obstet Gynecol*, 41(5): 491–499.
- Say, L., Chou, D., Gemmill, A., Tunçalp, Ö., Moller, A.-B., Daniels, J., Gülmezoglu, A. M., Temmerman, M. & Alkema, L. (2014). Global causes of maternal death: a WHO systematic analysis. *Lancet Glob Health*, 2(6): e323-333.

- Shen, M., Smith, G. N., Rodger, M., White, R. R., Walker, M. C. & Wen, S. W. (2017). Comparison of risk factors and outcomes of gestational hypertension and pre-eclampsia. *PLoS ONE*, 12(4): e0175914.
- Sildver, K., Veerus, P. & Lang, K. (2015). Sünnikaalukõverad Eestis ja sünnikaalu mõjutavad tegurid: registripõhine uuring. *Eesti Arst*, 94(8): 465–470.
- Staff, A. C. (2019). The two-stage placental model of preeclampsia: An update. *J Reprod Immunol*, 134–135: 1–10.
- Steinthorsdottir, V., McGinnis, R., Williams, N. O., ... Morgan, L. (2020). Genetic predisposition to hypertension is associated with preeclampsia in European and Central Asian women. *Nat Commun*, 11(1): 5976.
- Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G. & Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nat Rev Genet*, 20(8): 467–484.
- Tolcher, M. C., Sangi-Haghpeykar, H., Mendez-Figueroa, H. & Aagaard, K. M. (2020). Low-dose aspirin for preeclampsia prevention: efficacy by ethnicity and race. *Am J Obstet Gynecol* MFM, 2(4): 100184.
- Tranquilli, A. L., Brown, M. A., Zeeman, G. G., Dekker, G. & Sibai, B. M. (2013). The definition of severe and early-onset preeclampsia. Statements from the International Society for the Study of Hypertension in Pregnancy (ISSHP). *Pregnancy Hypertension: An International Journal of Women's Cardiovascular Health*, 3(1): 44–47.
- Tyrmi, J. S., Arffman, R. K., Pujol-Gualdo, N., ... Laivuori, H. (2022). Leveraging Northern European population history: novel low-frequency variants for polycystic ovary syndrome. *Hum Reprod*, 37(2): 352–365.
- Tyrmi, J. S., Kaartokallio, T., Lokki, A. I., ... FINNPEC Study Group, F. P., and the Estonian Biobank Research Team (2023). Genetic Risk Factors Associated With Preeclampsia and Hypertensive Disorders of Pregnancy. *JAMA Cardiology*, 8(7): 674–683.
- Uffelmann, E., Huang, Q. Q., Munung, N. S., de Vries, J., Okada, Y., Martin, A. R., Martin, H. C., Lappalainen, T. & Posthuma, D. (2021). Genome-wide association studies. *Nat Rev Methods Primers*, 1(1): 1–21.

- Umesawa, M. & Kobashi, G. (2017). Epidemiology of hypertensive disorders in pregnancy: prevalence, risk factors, predictors and prognosis. *Hypertens Res*, 40(3): 213–220.
- Veerbeek, J. H. W., Hermes, W., Breimer, A. Y., van Rijn, B. B., Koenen, S. V., Mol, B. W., Franx, A., de Groot, C. J. M. & Koster, M. P. H. (2015). Cardiovascular Disease Risk Factors After Early-Onset Preeclampsia, Late-Onset Preeclampsia, and Pregnancy-Induced Hypertension. *Hypertension*, 65(3): 600–606.
- Wang, K., Li, M. & Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*, 38(16): e164.
- Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. *Nat Commun*, 8(1): 1826.
- Welter, D., MacArthur, J., Morales, J., ... Parkinson, H. (2014). The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res*, 42(Database issue): D1001-1006.
- Weng, H., Li, Y., Fan, F., Yang, H., Zhou, G., Sun, P., Liu, S., Yi, T., Huo, Y. & Li, J. (2020). The association between total homocysteine and blood pressure in two independent Chinese populations. *J Hum Hypertens*, 34(9): 657–665.
- Wu, P., Haththotuwa, R., Kwok, C. S., ... Mamas, M. A. (2017). Preeclampsia and Future Cardiovascular Health: A Systematic Review and Meta-Analysis. *Circ Cardiovasc Qual Outcomes*, 10(2): e003497.
- Wu, P., Kwok, C. S., Haththotuwa, R., Kotronias, R. A., Babu, A., Fryer, A. A., Myint, P. K., Chew-Graham, C. A. & Mamas, M. A. (2016). Pre-eclampsia is associated with a twofold increase in diabetes: a systematic review and meta-analysis. *Diabetologia*, 59(12): 2518–2526.
- Xin, Q., Han, Y., Jiang, W., Wang, J., Luan, Y., Ji, Q. & Sun, W. (2022). Genetic susceptibility analysis of FGF5 polymorphism to preeclampsia in Chinese Han population. *Mol Genet Genomics*, 297(3): 791–800.
- Ying, W., Catov, J. M. & Ouyang, P. (2018). Hypertensive Disorders of Pregnancy and Future Maternal Cardiovascular Risk. *Journal of the American Heart Association*, 7(17): e009382.

Kasutatud veebiaadressid

CTG–VL veebiplatvorm: <https://vl.genoma.io/> (26.10.2022)

FinnGen andmebaasi dokumentatsioon: https://www.finnngen.fi/en/access_results (19.08.2023)

FinnGen R7 andmebaas: <http://r7.finnngen.fi> (20.08.2022)

FUMA veebiplatvorm: <https://fuma.ctglab.nl/> (20.08.2023)

Illumina koduleht: <https://www.illumina.com/> (19.09.2023)

LD-skoori regressiooni kasutusjuhend: <https://github.com/bulik/ldsc> (09.09.2022)

Rahvusvahelise haiguste klassifikatsiooni 10. väljaanne: <https://icd.who.int/browse10/2019/en>
(11.09.2022)

R tarkvara: <https://www.r-project.org/> (30.08.2022)

Soome-spetsiifiline referentspaneel SISu: <http://www.sisuproject.fi/> (19.08.2023)

Tervise Arengu Instituudi tervisestatistika ja terviseuuringute andmebaas:
<https://statistika.tai.ee/> (26.04.2023)

Vaadeldud päritavuse ja selle SD teisaldamise valem:

<https://gist.github.com/nievergeltlab/fb8a20feded72030907a9b4e81d1c6ea> (27.09.2022)

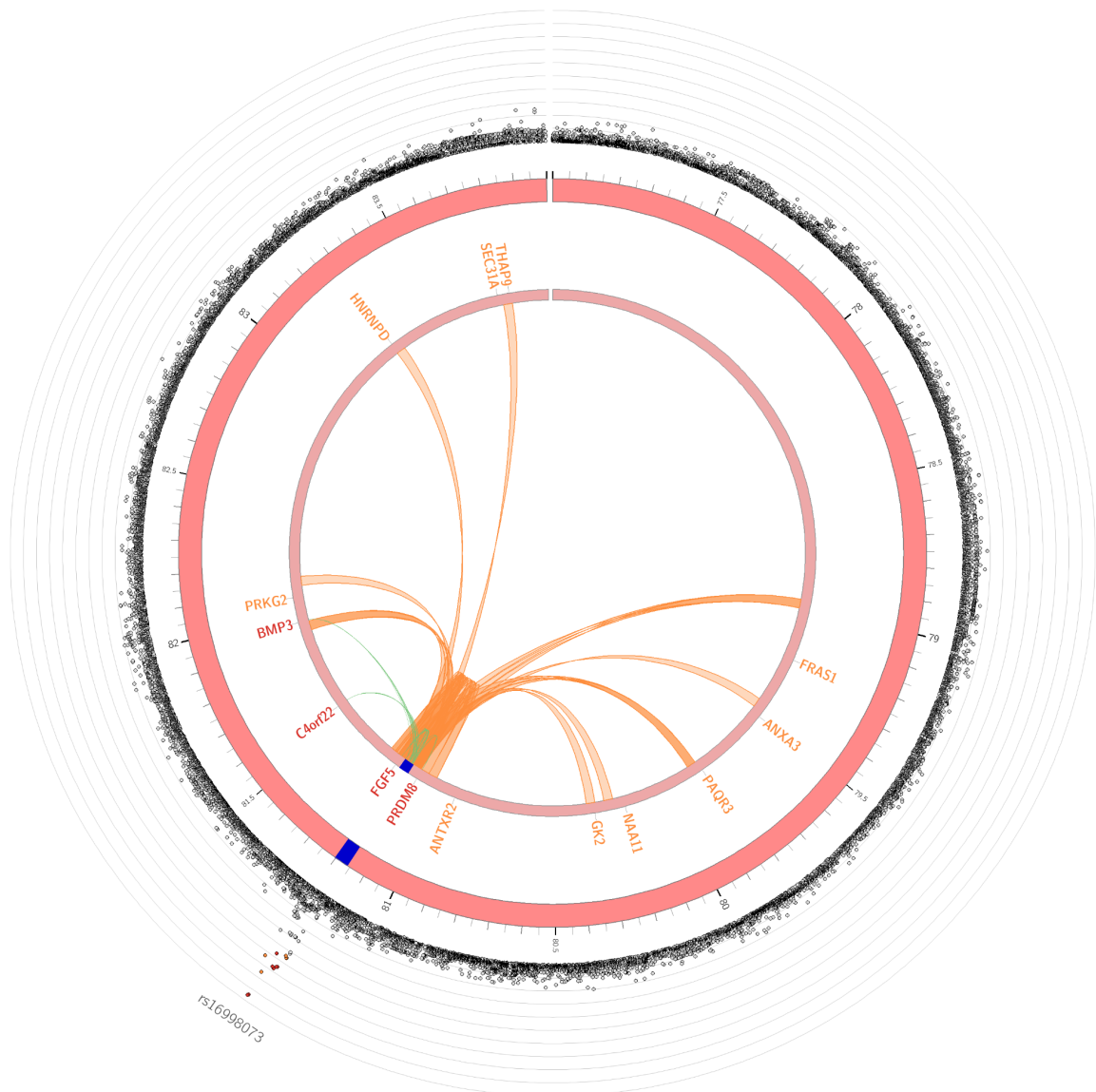
Lisad

Lisa 1.

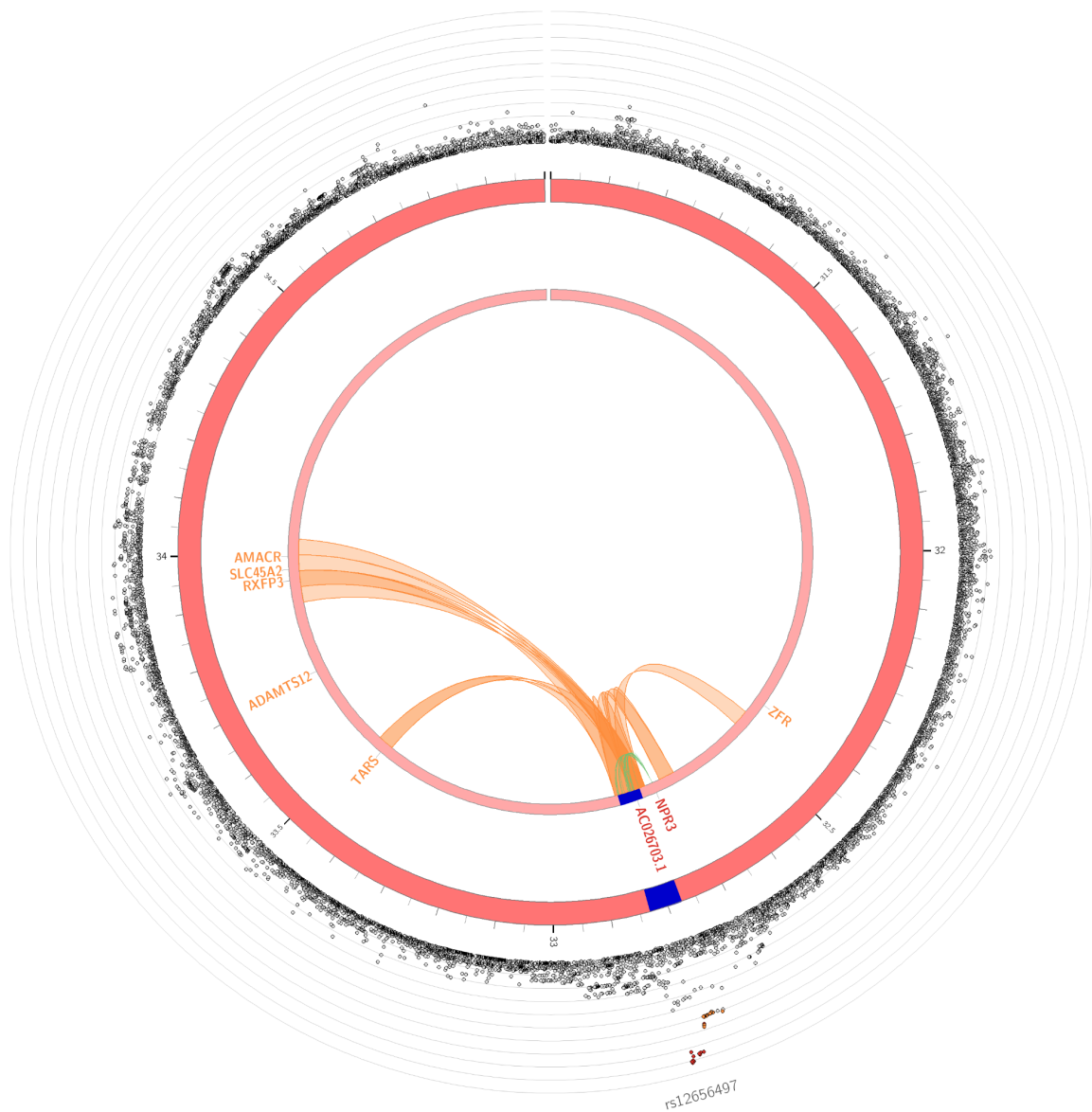
Tabel 7. Rasedushüpertensiooni juhtivate SNP-de p-väärtused ja OR (95% CI) preeklampsia GWA metaanalüüsis. OR – šansside suhe (*odds ratio*); 95% CI – šansside suhte 95%-usaldusintervall (*confidence interval*). Paksus kirjas on ära toodud *subsignificant* p-väärtused.

Marker	p-väärtus	OR (95% CI)
rs143439093	$1,44 \times 10^{-3}$	1,13 (1,05–1,21)
rs16998073	$1,39 \times 10^{-10}$	0,88 (0,85–0,92)
rs12656497	$3,57 \times 10^{-4}$	0,94 (0,90–0,97)
rs7139122	$7,15 \times 10^{-4}$	0,81 (0,71–0,91)
rs167479	$4,20 \times 10^{-7}$	1,10 (1,06–1,14)
rs2208589	$9,17 \times 10^{-4}$	0,91 (0,86–0,96)
rs74715624	$6,17 \times 10^{-8}$	0,88 (0,84–0,92)

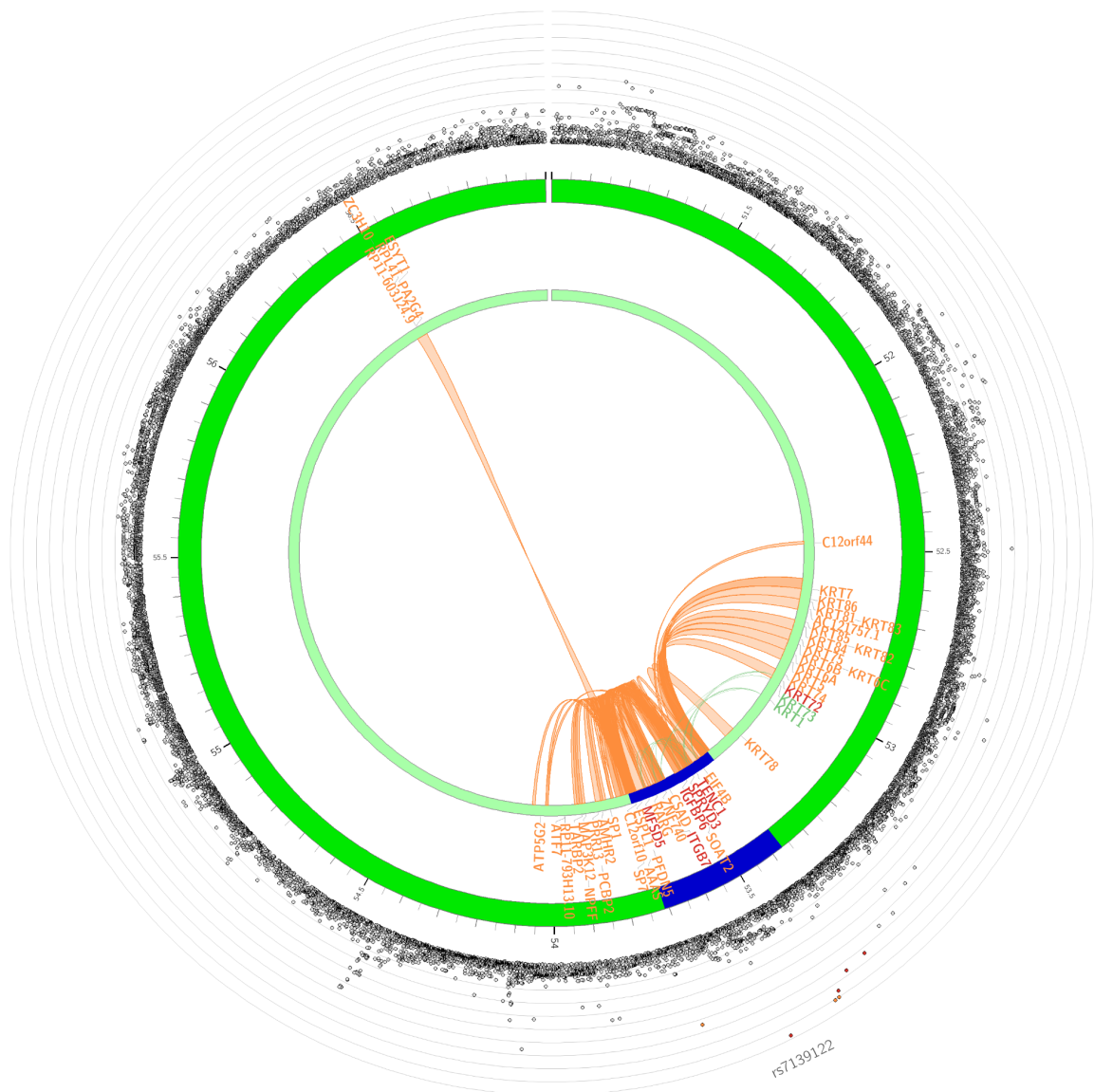
Lisa 2.



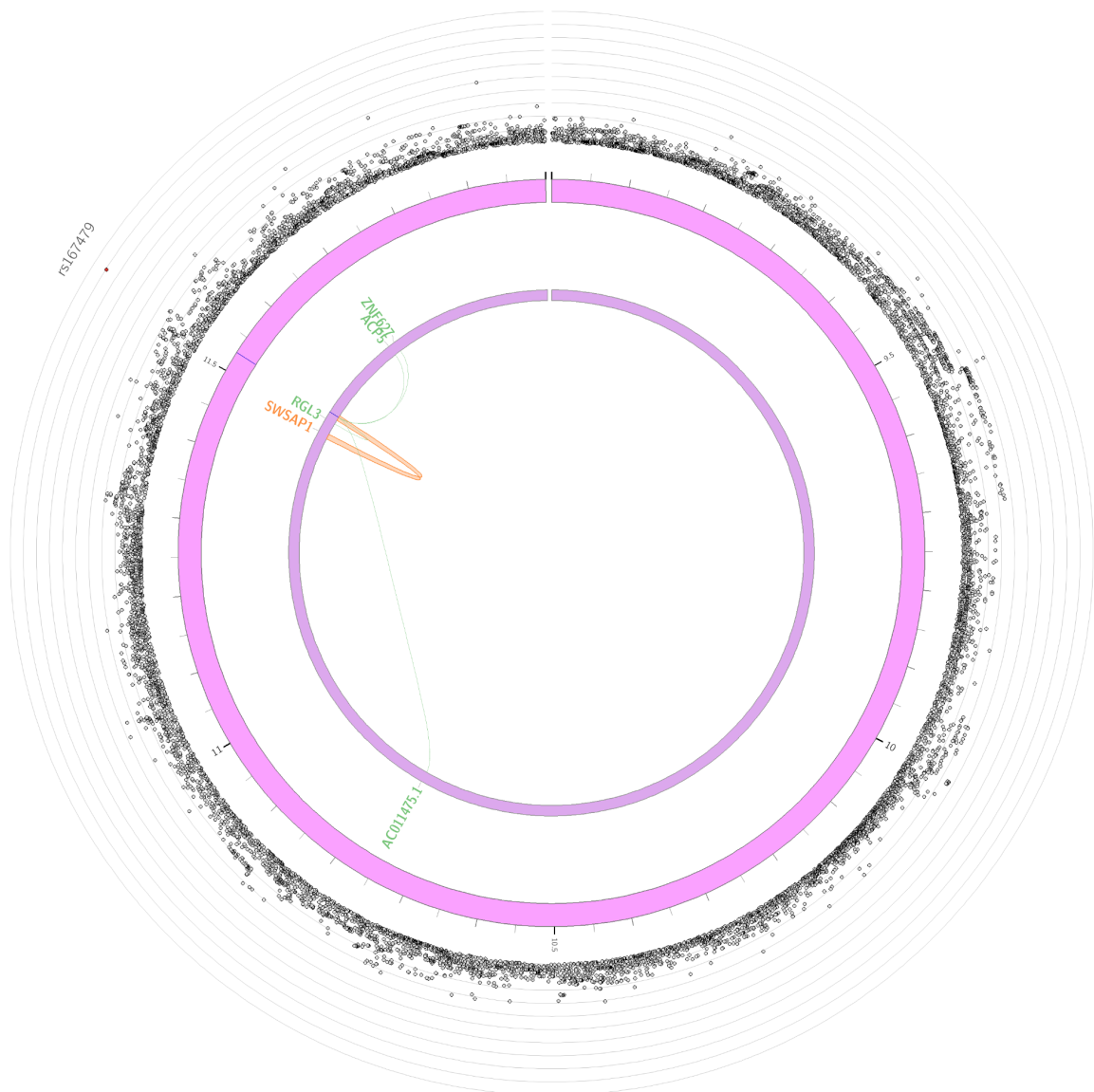
Joonis 10. Rasedushüpertensiooni GWAS metaanalüüsi tulemuste põhjal valmistatud 4. kromosoomi *circos*'e joonis. Välimisele kihile on kantud Manhattani joonis. Joonisele on kantud geneetilised variandid, mille p -väärtus $\leq 0,05$. Y-teljel on vastava geneetilise variandi p -väärtuse negatiivne kümnendlogaritm. Juhtiva SNP-i unikaalne kood on hallis kirjas. SNP-d on tähistatud vastavalt nende suurimale r^2 väärtusele ühega riskilookuses olevatest sõltumatutest olulistest SNP-dest: punane ($r^2 > 0,8$), oranž ($r^2 > 0,6$), roheline ($r^2 > 0,4$), sinine ($r^2 > 0,2$) ja hall ($r^2 \leq 0,2$). Keskmisele kihile on kantud kromosoom koos koordinaatidega, sisemisele kihile – koordinaatideta; sinise värviga on tähistatud riskilookus. Kandidaatgeenid on tähistatud vastavalt toetusandmete olemusele: roheline (eQTL), oranž (kromatiininteraktsioon) ja punane (eQTL ja kromatiininteraktsioon). Sisemise kihi kaared tähistavad eQTL-e (roheline) ja kromatiininteraktsioone (oranž). Joonis on tehtud FUMA veebiplatvormi abil.



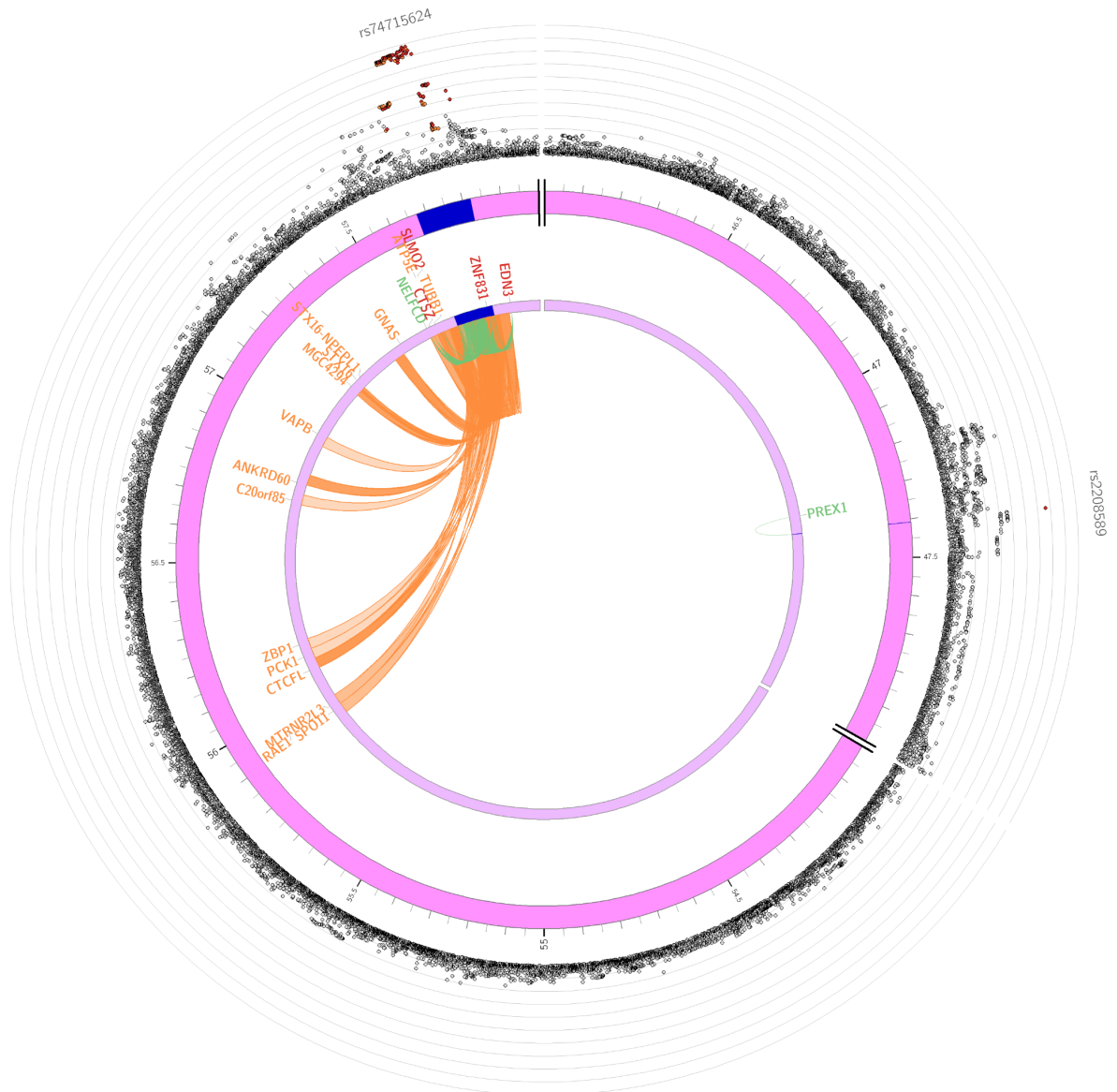
Joonis 11. Rasedushüpertensiooni GWAS metaanalüüsi tulemuste põhjal valmistatud 5. kromosoomi *circos*'e joonis. Välimisele kihile on kantud Manhattani joonis. Joonisele on kantud geneetilised variandid, mille p -väärtus $\leq 0,05$. Y-teljel on vastava geneetilise variandi p -väärtuse negatiivne kümnendlogaritm. Juhtiva SNP-i unikaalne kood on hallis kirjas. SNP-d on tähistatud vastavalt nende suurimale r^2 väärtusele ühega riskilookuses olevatest sõltumatutest olulistest SNP-dest: punane ($r^2 > 0,8$), oranž ($r^2 > 0,6$), roheline ($r^2 > 0,4$), sinine ($r^2 > 0,2$) ja hall ($r^2 \leq 0,2$). Keskmisele kihile on kantud kromosoom koos koordinaatidega, sisemisele kihile – koordinaatideta; sinise värviga on tähistatud riskilookus. Kandidaatgeenid on tähistatud vastavalt toetusandmete olemusele: roheline (eQTL), oranž (kromatiininteraktsioon) ja punane (eQTL ja kromatiininteraktsioon). Sisemise kihi kaared tähistavad eQTL-e (roheline) ja kromatiininteraktsioone (oranž). Joonis on tehtud FUMA veebiplatvormi abil.



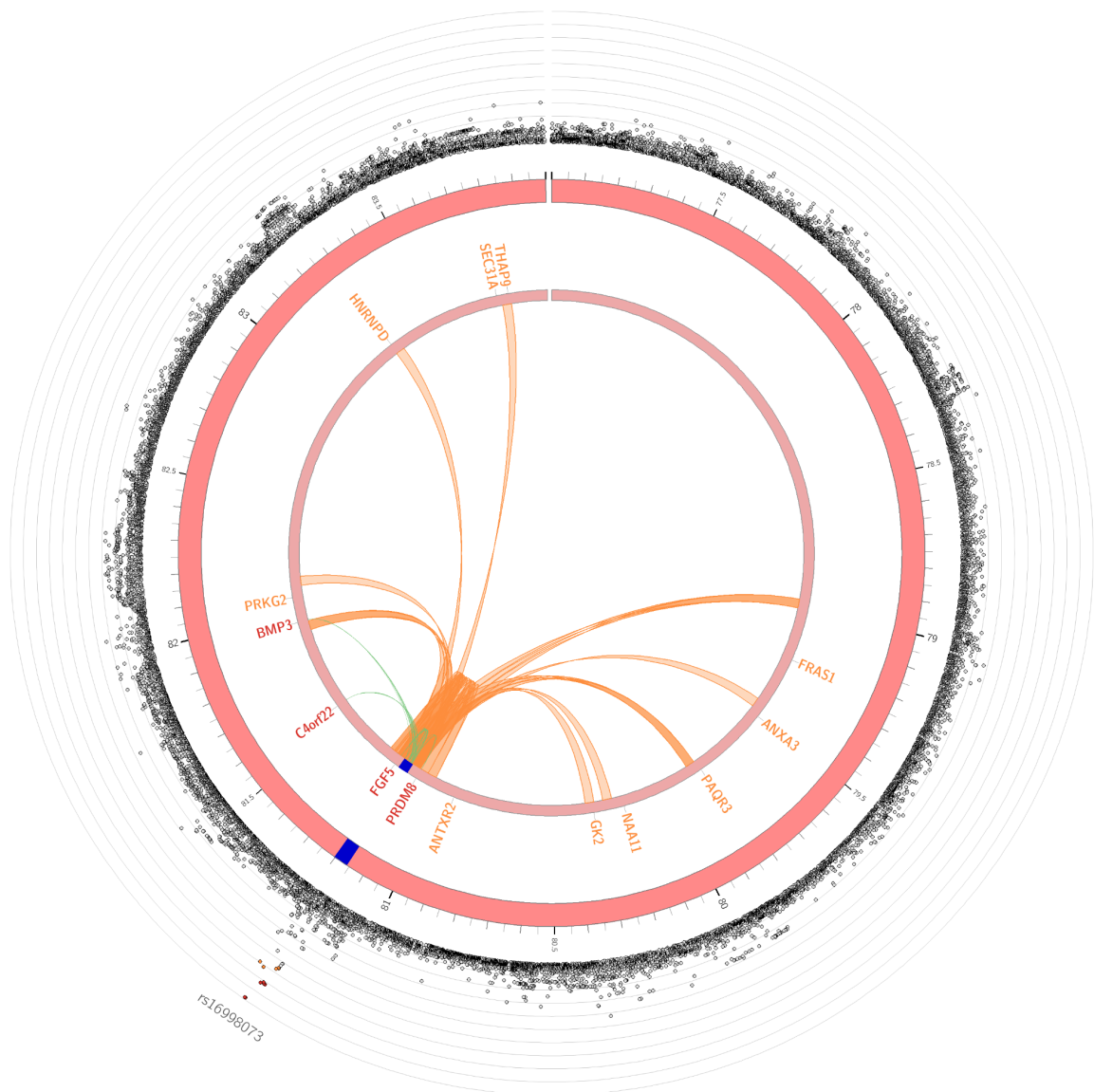
Joonis 12. Rasedushüpertensiooni GWAS metaanalüüsi tulemuste põhjal valmistatud 12. kromosoomi *circos*'e joonis. Välimisele kihile on kantud Manhattani joonis. Joonisele on kantud geneetilised variandid, mille p -väärtus $\leq 0,05$. Y-teljel on vastava geneetilise variandi p -väärtuse negatiivne kümnendlogaritm. Juhtiva SNP-i unikaalne kood on hallis kirjas. SNP-d on tähistatud vastavalt nende suurimale r^2 väärtusele ühega riskilookuses olevatest sõltumatutest olulistest SNP-dest: punane ($r^2 > 0,8$), oranž ($r^2 > 0,6$), roheline ($r^2 > 0,4$), sinine ($r^2 > 0,2$) ja hall ($r^2 \leq 0,2$). Keskmisele kihile on kantud kromosoom koos koordinaatidega, sisemisele kihile – koordinaatideta; sinise värviga on tähistatud riskilookus. Kandidaatgeenid on tähistatud vastavalt toetusandmete olemusele: roheline (eQTL), oranž (kromatiininteraktsioon) ja punane (eQTL ja kromatiininteraktsioon). Sisemise kihi kaared tähistavad eQTL-e (roheline) ja kromatiininteraktsioone (oranž). Joonis on tehtud FUMA veebiplatvormi abil.



Joonis 13. Rasedushüpertensiooni GWAS metaanalüüsi tulemuste põhjal valmistatud 19. kromosoomi *circos*'e joonis. Välimisele kihile on kantud Manhattani joonis. Joonisele on kantud geneetilised variandid, mille p -väärtus $\leq 0,05$. Y-teljel on vastava geneetilise variandi p -väärtuse negatiivne kümnendlogaritm. Juhtiva SNP-i unikaalne kood on hallis kirjas. SNP-d on tähistatud vastavalt nende suurimale r^2 väärtusele ühega riskilookuses olevatest sõltumatutest olulistest SNP-dest: punane ($r^2 > 0,8$), oranž ($r^2 > 0,6$), roheline ($r^2 > 0,4$), sinine ($r^2 > 0,2$) ja hall ($r^2 \leq 0,2$). Keskmisele kihile on kantud kromosoom koos koordinaatidega, sisemisele kihile – koordinaatideta; sinise värviga on tähistatud riskilookus. Kandidaatgeenid on tähistatud vastavalt toetusandmete olemusele: roheline (eQTL), oranž (kromatiininteraktsioon) ja punane (eQTL ja kromatiininteraktsioon). Sisemise kihi kaared tähistavad eQTL-e (roheline) ja kromatiininteraktsioone (oranž). Joonis on tehtud FUMA veebiplatvormi abil.

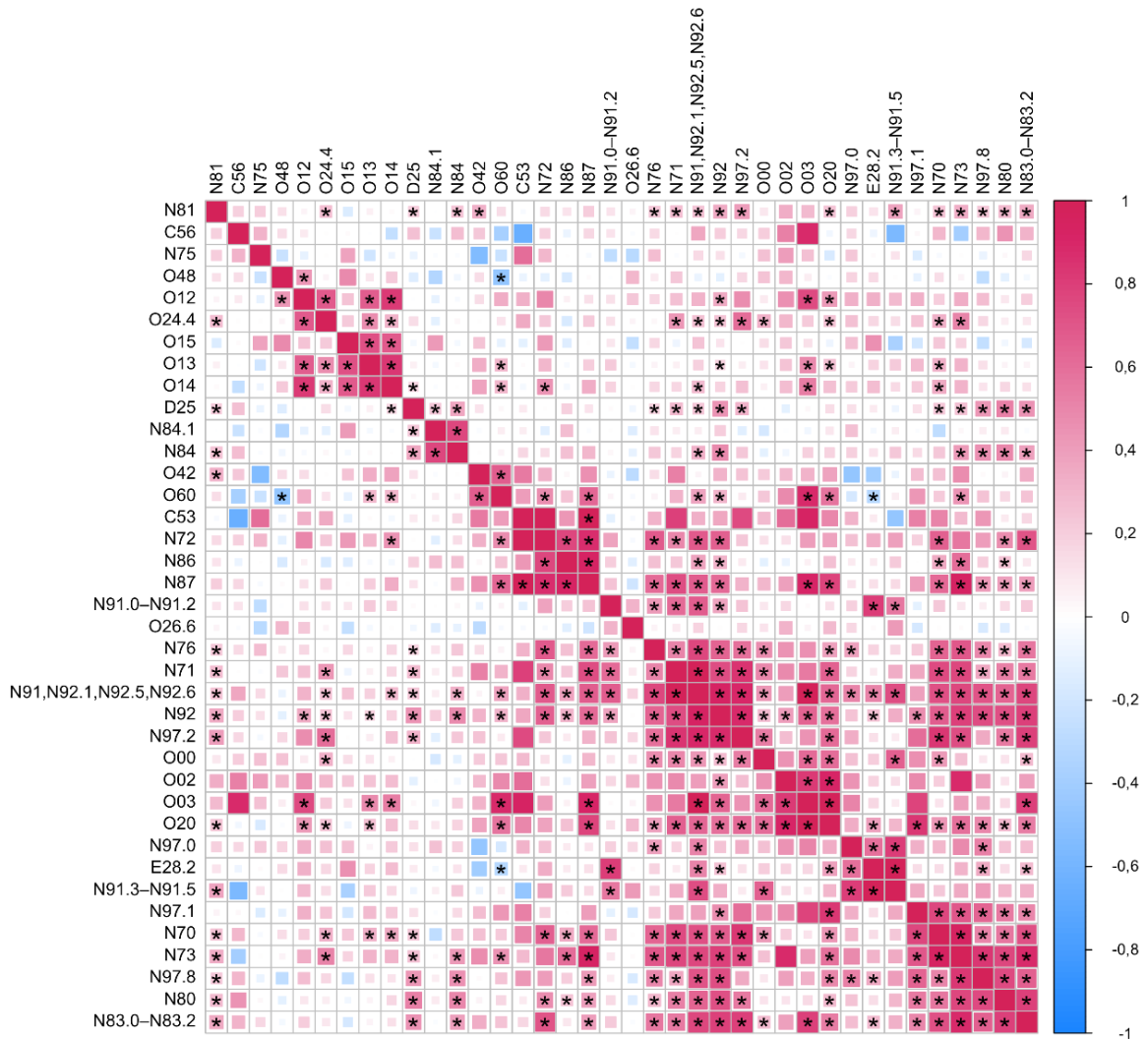


Joonis 14. Rasedushüpertensiooni GWAS metaanalüüsi tulemuste põhjal valmistatud 20. kromosoomi *circos*'e joonis. Välimisele kihile on kantud Manhattani joonis. Joonisele on kantud geneetilised variandid, mille p -väärtus $\leq 0,05$. Y-teljel on vastava geneetilise variandi p -väärtuse negatiivne kümnendlogaritm. Juhtiva SNP-i unikaalne kood on hallis kirjas. SNP-d on tähistatud vastavalt nende suurimale r^2 väärtusele ühega riskilookuses olevatest sõltumatutest olulistest SNP-dest: punane ($r^2 > 0,8$), oranž ($r^2 > 0,6$), roheline ($r^2 > 0,4$), sinine ($r^2 > 0,2$) ja hall ($r^2 \leq 0,2$). Keskmisele kihile on kantud kromosoom koos koordinaatidega, sisemisele kihile – koordinaatideta; sinise värviga on tähistatud riskilookus. Kandidaatgeenid on tähistatud vastavalt toetusandmete olemusele: roheline (eQTL), oranž (kromatiininteraktsioon) ja punane (eQTL ja kromatiininteraktsioon). Sisemise kihi kaared tähistavad eQTL-e (roheline) ja kromatiininteraktsioone (oranž). Joonis on tehtud FUMA veebiplatvormi abil.



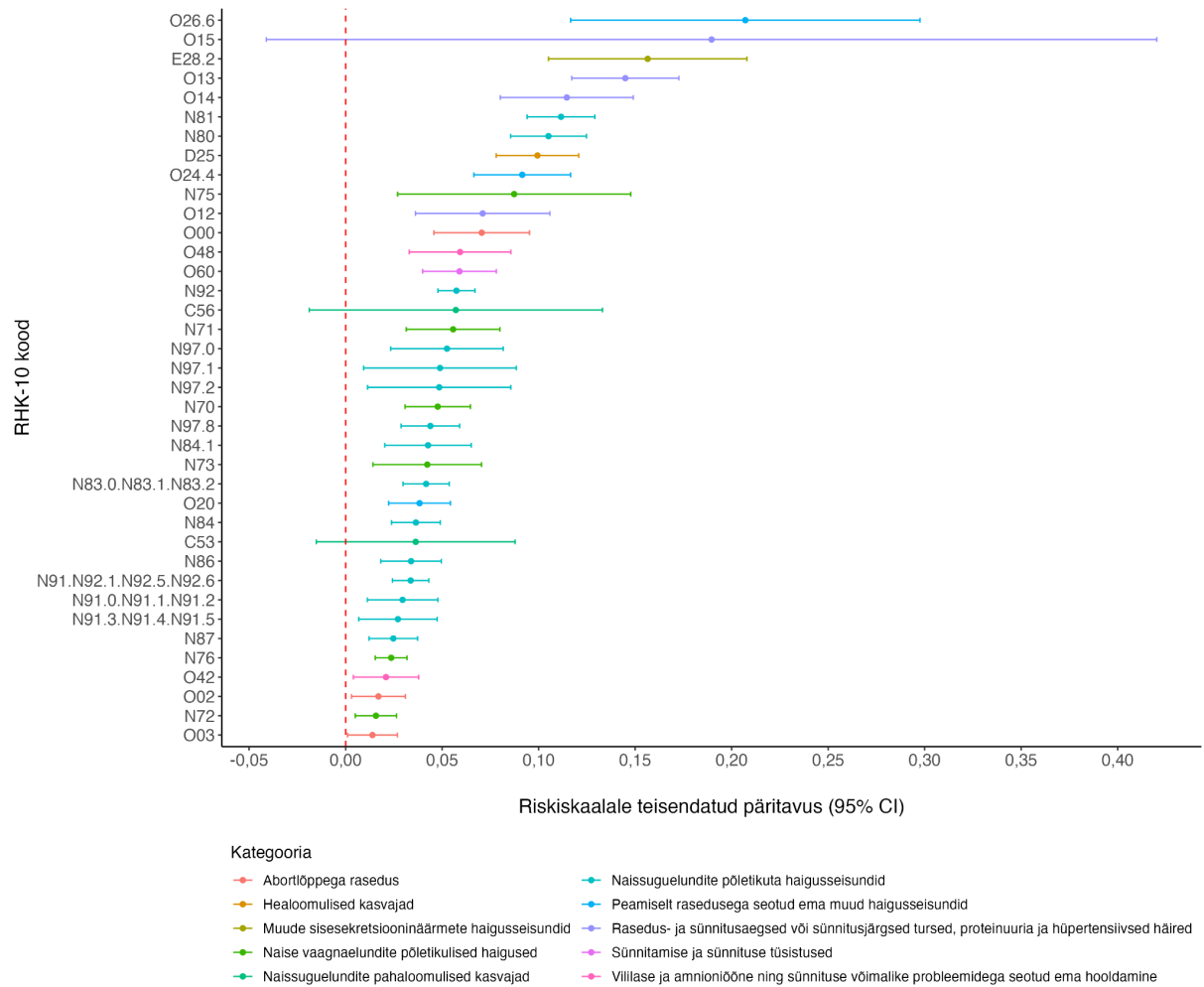
Joonis 15. Preeklampsia GWAS metaanalüüsi tulemuste põhjal valmistatud 4. kromosoomi *circos*'e joonis. Välimisele kihile on kantud Manhattani joonis. Joonisele on kantud geneetilised variandid, mille p -väärtus $\leq 0,05$. Y-teljel on vastava geneetilise variandi p -väärtuse negatiivne kümnendlogaritm. Juhtiva SNP-i unikaalne kood on hallis kirjas. SNP-d on tähistatud vastavalt nende suurimale r^2 väärtusele ühega riskilookuses olevatest sõltumatutest olulistest SNP-dest: punane ($r^2 > 0,8$), oranž ($r^2 > 0,6$), roheline ($r^2 > 0,4$), sinine ($r^2 > 0,2$) ja hall ($r^2 \leq 0,2$). Keskmisele kihile on kantud kromosoom koos koordinaatidega, sisemisele kihile – koordinaatideta; sinise värviga on tähistatud riskilookus. Kandidaatgeenid on tähistatud vastavalt toetusandmete olemusele: roheline (eQTL), oranž (kromatiininteraktsioon) ja punane (eQTL ja kromatiininteraktsioon). Sisemise kihi kaared tähistavad eQTL-e (roheline) ja kromatiininteraktsioone (oranž). Joonis on tehtud FUMA veebiplatvormi abil.

Lisa 3.



Joonis 15. 38 uuritava tunnuse geneetiliste korrelatsioonide sümmeetriline maatriks. Roosa värviga on tähistatud positiivsed seosed, sinise värviga on tähistatud negatiivsed seosed. Ruudu suurus peegeldab r_g absoluutväärtust. Diagnoosid on esitatud RHK-10 tähtnumbriliste koodide kujul. Tärniga („*“) on tähistatud statistiliselt olulised (korrigeeritud p-väärtus < 0,05) seosed. RHK-10 koodide vasted on toodud välja Lisis 5.

Lisa 4.



Joonis 16. Riskiskaalale teisaldatud päritavus ($h_{SNP L}^2$) 38 raseduse ja naiste tervisega seotud diagnoosi jaoks. Diagnoosid on esitatud RSK-10 tähtnumbriliste koodide kujul. Diagnoosi värv vastab laiemale kategooriale, kuhu antud tunnus kuulub. Hinnanguline SNP-põhine päritavus on kantud joonisele punktina. Veajoon tähistab 95%-usaldusintervalli. Diagnoosid on järjestatud päritavuskoeffitsiendi väärtuse kahanemise järjekorras. Punane katkendjoon näitab päritavuse komponendi puudumist. RSK-10 koodide vasted on toodud välja Lisas 5.

Lisa 5.

Tabel 8. RHK-10 tähtnumbriliste koodide kirjeldus.

RHK-10 kood	Nimetus
C53	Emakakaela pahaloomuline kasvaja
C56	Munasarja pahaloomuline kasvaja
D25	Emaka leiomüoom
E28.2	Polütsüstilise munasarja sündroom
N70	Salpingiit e munajuhapõletik ja ooforiit e munasarjapõletik
N71	Emaka (v.a emakakael) põletikuline haigus
N72	Emakakaela põletikuline haigus
N73	Naise vaagnapiirkonna muud põletikulised haigused
N75	Bartholini näärme e suure esikunäärme haigused
N76	Tupe ja häbeme muu põletik
N80	Endometrioos
N81	Naissuguelundite prolaps e väljalange
N83.0–N83.2	Munasarjatsüstid
N84	Naissuguelundite polüüp
N84.1	Emakakaelapolüüp
N86	Emakakaela erosioon e limaskestamarrastus ja ektroopion e väljapööre
N87	Emakakaela düsplaasia
N91,N92.1,N92.5,N92.6	Korratu menstruatsioon
N91.0–N91.2	Amenorröa e menstruatsioonipuudus
N91.3–N91.5	Oligomenorröa e harv menstruatsioon
N92	Rohke, sage ja korratu menstruatsioon
N97.0	Ovulatsioonipuudusega kaasnev naiseviljatus
N97.1	Munajuha[de]tekkene naiseviljatus
N97.2	Emakatekkene naiseviljatus
N97.8	Naiseviljatus muul põhjusel
O00	Emakaväline e ektoopiline rasedus
O02	Eostumise muud anomaalsed tulemused
O03	Spontaanabort e iseeneslik abort
O12	Hüpertensioonita rasedusturse ja proteinuuria
O13	Olulise proteinuuriata rasedushüpertensioon
O14	Olulise proteinuuriaga rasedushüpertensioon

O15	Eklampsia e kramptõbi
O20	Varane rasedusaegne verejooks
O24.4	Rasedusajal tekkiv suhkurtõbi
O26.6	Maksa haigusseisundid raseduse, sünnitamise ja sünnitusjärgse perioodi ajal
O42	Lootevee enneaegne puhkemine
O48	Ülekantud rasedus
O60	Enneaegne sünnitus

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Valentina Rukins,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose
“Preeklampsia ja rasedusaegse hüpertensiooni geneetika ning seos teiste diagnoosidega TÜ Eesti Geenivaramu ja FinnGen andmetel”,
mille juhendajad on kaasprofessor Triin Laisk ja professor Reedik Mägi,
reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.
2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commonsi litsentsiga CC BY NC ND 4.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.
3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.
4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Valentina Rukins

21.08.2023