



Egija Berga-Švītiņa

Ģenētisko faktoru, kas saistīti ar krūts vai olnīcu vēža risku, identificēšana *BRCA1* patogēno variantu nesējās

Promocijas darba kopsavilkums zinātnes doktora grāda
“zinātnes doktors (*Ph. D.*)” iegūšanai

Nozaru grupa – medicīnas un veselības zinātnes
Nozare – medicīnas bāzes zinātnes, tai skaitā farmācija
Apakšnozare – medicīniskā ģenētika

Rīga, 2024



RĪGAS STRADIŅA
UNIVERSITĀTE

Egija Berga-Švītiņa

ORCID 0000-0001-5150-0185

Ģenētisko faktoru, kas saistīti ar krūts
vai olnīcu vēža risku, identificēšana
BRCA1 patogēno variantu nesējās

Promocijas darba kopsavilkums zinātnes doktora grāda
“zinātnes doktors (*Ph. D.*)” iegūšanai

Nozaru grupa – medicīnas un veselības zinātnes

Nozare – medicīnas bāzes zinātnes, tai skaitā farmācija

Apakšnozare – medicīniskā ģenētika

Rīga, 2024

Promocijas darbs izstrādāts Rīgas Stradiņa universitātes Onkoloģijas un molekulārās ģenētikas institūtā, Latvijā

Promocijas darba vadītāji:

Dr. biol. profesors **Edvīns Miklaševičs**, Rīgas Stradiņa universitātes
Onkoloģijas un molekulārās ģenētikas institūts;
Bioloģijas un mikrobioloģijas katedra, Latvija

Dr. rer. nat. **Baiba Vilne**, Rīgas Stradiņa universitātes
Zinātnes centra Bioinformātikas grupa, Latvija

Zinātniskā konsultante:

Dr. med. **Jeļena Maksimenko**, Rīgas Stradiņa universitātes
Onkoloģijas un molekulārās ģenētikas institūts; Paula Stradiņa Klīniskās
universitātes slimnīcas Krūts slimību centra Krūts ķirurģijas nodaļa, Latvija

Oficiālie recenzenti:

Dr. med. asociētā profesore **Madara Auzenbaha**, Rīgas Stradiņa
universitātes Bioloģijas un mikrobioloģijas katedra, Latvija

PhD **Raitis Pečulis**, Latvijas Biomedicīnas pētījumu un studiju centrs, Latvija

MD, PhD profesors **Neeme Tõnisson**, Igaunijas Biobanka;
Tartu universitātes Genomikas institūts, Igaunija

Promocijas darbs tiks aizstāvēts Medicīnas bāzes zinātņu, tai skaitā farmācijas, promocijas padomes atklātā sēdē 2024. gada 20. septembrī plkst. 10.00 attālināti, tiešsaistes platformā *Zoom*.

Ar promocijas darbu var iepazīties RSU bibliotēkā un RSU tīmekļa vietnē:
<https://www.rsu.lv/promocijas-darbi>

NACIONĀLAIS
ATTĪSTĪBAS
PLANS 2020



EUROPAS SAVIENĪBA
Eiropas Sociālais
fonds

IEGULDĪJUMS TAVĀ NĀKOTNĒ

Pētījums veikts projekta Nr. 8.2.2.0/20/I/004 “Atbalsts doktorantu iesaistei zinātniski pētnieciskajā un studiju darbā” ietvaros

Promocijas padomes sekretārs:

Dr. sc. soc. asociētais profesors **Andrejs Ivanovs**

Satura rādītājs

Darbā izmantotie saīsinājumi.....	5
Ievads.....	8
Darba mērķis.....	9
Darba uzdevumi.....	10
Darba hipotēze.....	10
Darba novitāte.....	10
1. Materiāli un metodes.....	11
1.1. Pētījuma kohorta.....	11
1.2. <i>BRCA1</i> un <i>CHEK2</i> dubultheterozigotu analīze.....	12
1.3. Paraugu genotipēšana ar <i>OncoArray-500K</i> hromosomu mikrorindu analīzes metodi.....	13
1.4. Genotipu noteikšana un kvalitātes kontrole (<i>QC</i>).....	14
1.5. Genotipu imputācija.....	15
1.6. Genoma mēroga asociāciju analīze (<i>GWAS</i>), izmantojot <i>SAIGE</i>	16
1.7. Pēc- <i>GWAS</i> analīze, izmantojot brīvpiekļuves platformas <i>FUMA</i> un <i>VEP</i>	16
1.8. Poligēnā riska modeļa (PRS) aprēķini.....	17
1.9. Statistiskā analīze.....	18
1.10. Datu pieejamība.....	19
2. Rezultāti.....	20
2.1. Pētījuma kohortas raksturojums.....	20
2.1.1. Pacientu raksturojums.....	20
2.1.2. <i>BRCA1</i> PV c.4035del un c.5266dup penetrance pētījuma kohortā.....	21
2.1.3. Ar vecumu saistītā KV vai OV kumulatīvā incidence <i>BRCA1</i> c.4035del un c.5266dup PV nesējās.....	22
2.2. Pētījuma dizains: uz hipotēzi balstīta vs. uz datiem balstīta analīze.....	25
2.2.1. Uz hipotēzi balstīta <i>BRCA1</i> un <i>CHEK2</i> dubultheterozigotu analīze.....	25
2.2.2. Uz datiem balstīta ar vēža risku saistītu variantu identificēšana viena nukleotīda varianta (<i>SNV</i>) līmenī <i>BRCA1</i> PV nesējās.....	27
2.2.3. Uz datiem balstīta ar vēža risku saistītu variantu identificēšana poligēnu riska variantu (PRS) līmenī <i>BRCA1</i> PV nesējās.....	32

3. Diskusija	35
3.1. Galvenie rezultāti pētījuma kohortā.....	35
3.1.1. Uz hipotēzi balstīta <i>BRCA1</i> un <i>CHEK2</i> dubultheterozigotu analīze	37
3.1.2. Uz datiem balstīta ar vēža risku saistītu variantu identificēšana viena nukleotīda varianta (<i>SNV</i>) līmenī <i>BRCA1</i> PV nesējās.....	38
3.1.3. Uz datiem balstīta ar vēža risku saistītu variantu identificēšana poligēnu riska variantu (<i>PRS</i>) līmenī <i>BRCA1</i> PV nesējās.....	44
3.2. Pētījuma priekšrocības un ierobežojumi	47
3.3. Nākotnes perspektīvas	48
Secinājumi	50
Priekšlikumi	51
Publikāciju, ziņojumu un patentu saraksts par promocijas darba tēmu.....	52
Literatūras un avotu saraksts	53
Pateicības	62

Darbā izmantotie saīsinājumi

Saīsinājums	Skaidrojums latviešu valodā	Skaidrojums angļu valodā
<i>ACMG</i>	Amerikas medicīniskās ģenētikas un genomikas koledža	<i>American College of Medical Genetics and Genomics</i>
<i>AIM</i>	informatīvs priekšteču marķieris	<i>ancestry informative marker</i>
<i>ALT</i>	alternatīvā alēle	<i>alternative allele</i>
<i>ARL3</i>	ADP ribozilēšanas faktoram līdzīgā GTPāze 3	<i>ADP ribosylation factor like GTPase 3</i>
<i>AUC</i>	laukums zem līknes	<i>area under the curve</i>
<i>BCAC</i>	krūts vēža asociācijas konsorcijs	<i>Breast Cancer Association Consortium</i>
<i>BRCA1</i>	ar <i>BRCA1</i> DNS reparāciju saistītais	<i>BRCA1 DNA repair associated</i>
<i>BRCA2</i>	ar <i>BRCA2</i> DNS reparāciju saistītais	<i>BRCA2 DNA repair associated</i>
<i>BRRM</i>	abpusēja risku samazinoša mastektomija	<i>bilateral risk-reducing mastectomy</i>
<i>BRRSO</i>	abpusēja risku samazinoša salpingooforektomija	<i>bilateral risk-reducing salpingo-oophorectomy</i>
<i>CDPC</i>	slimību profilakses un kontroles centrs	<i>Centre for Disease Prevention and Control</i>
<i>CHEK2</i>	kontrolpunkta kināze 2	<i>checkpoint kinase 2</i>
<i>CIMBA</i>	<i>BRCA1/2</i> modificējošo faktoru pētnieku konsorcijs	<i>Consortium of Investigators of Modifiers of BRCA1/2</i>
<i>DNS</i>	dezoksiribonukleīnskābe	<i>deoxyribonucleic acid</i>
<i>DNAAF9</i>	dineīna aksonemālās montāžas faktors 9	<i>dynein axonemal assembly factor 9</i>
<i>DR2</i>	devas R-kvadrāts	<i>dosage R-squared</i>
<i>eQTL</i>	ekspresijas kvantitatīvo pazīmju lokusi	<i>expression quantitative trait loci</i>

<i>FAM107B</i>	saimes ar sekvenses līdzību 107 loceklis B	<i>family with sequence similarity 107 member B</i>
<i>FDR</i>	viltus atklāšanas kļūda	<i>false discovery rate</i>
<i>FUMA</i>	funkcionālā kartēšana un anotācija	<i>functional mapping and annotation</i>
<i>GLOBOCAN</i>	Globālā vēža observatorija	<i>global cancer observatory</i>
<i>GRCh37</i>	genoma references konsorcijs cilvēka genoms 37	<i>genome reference consortium human build 37</i>
<i>GTE_x</i>	genotipa–audu ekspresijas projekts	<i>genotype-tissue expression project</i>
<i>GWAS</i>	genoma mēroga asociāciju analīze	<i>genome-wide association study</i>
<i>HBOC</i>	pārmantots krūts un olnīcu vēzis	<i>hereditary breast and ovarian cancer</i>
<i>HR</i>	riska attiecība	<i>hazard ratio</i>
<i>HWE</i>	Hārdija–Veinberga likums	<i>Hardy–Weinberg equilibrium</i>
<i>IBD</i>	identitāte pēc izcelsmes	<i>identity by descent</i>
<i>KV</i>	krūts vēzis	<i>breast cancer</i>
<i>LD</i>	nelīdzsvarotā saistība	<i>linkage disequilibrium</i>
<i>LGDB</i>	Latvijas iedzīvotāju genomu datubāze	<i>genome database of Latvian population</i>
<i>lncRNA</i>	garā nekodējošā RNS	<i>long non-coding RNA</i>
<i>LoF</i>	funkcijas zudums	<i>loss-of-function</i>
<i>MAF</i>	retās alēles biežums	<i>minor allele frequency</i>
<i>NES</i>	normalizētais efekta lielums	<i>normalized effect size</i>
<i>OCAC</i>	olnīcu vēža asociācijas konsorcijs	<i>Ovarian Cancer Association Consortium</i>
<i>OR</i>	izredžu attiecība	<i>odds ratio</i>
<i>OV</i>	olnīcu vēzis	<i>ovarian cancer</i>

PC	principālā komponente	<i>principal component</i>
PCR	polimerāzes ķēdes reakcija	<i>polymerase chain reaction</i>
PRS	poligēnā riska modelis	<i>polygenic risk score</i>
PV	patogēns variants	<i>pathogenic variant</i>
QC	kvalitātes kontrole	<i>quality control</i>
REF	references alēle	<i>reference allele</i>
RNS	ribonukleīnskābe	<i>ribonucleic acid</i>
ROC	uztvērēja darbības raksturlīkne	<i>receiver-operating characteristic</i>
SD	standartnovirze	<i>standard deviation</i>
SE	standartkļūda	<i>standard error</i>
SLC1A5	izšķīdušo vielu nesēju 1. saimes 5. loceklis	<i>solute carrier family 1 member 5</i>
SNV	viena nukleotīda variants	<i>single nucleotide variant</i>
TI	ticamības intervāls	<i>confidence interval</i>
TNBC	trīskārši negatīvs krūts vēzis	<i>triple-negative breast cancer</i>
VEP	varianta efekta prognozētājs	<i>variant effect predictor</i>
WES	pilna eksoma sekvencēšana	<i>whole exome sequencing</i>
WGS	pilna genoma sekvencēšana	<i>whole genome sequencing</i>
ZNF514	cinka pirkstu proteīns 514	<i>zinc finger protein 514</i>

Ievads

Saskaņā ar Globālās vēža observatorijas (*GLOBOCAN*) 2020. gada datiem krūts vēzis (KV) ir visbiežāk diagnosticētais vēzis sievietēm, kas izraisa 15 % ar vēzi saistīto nāves gadījumu visā pasaulē, ar aptuveni 522 000 reģistrētiem gadījumiem. Savukārt olnīcu vēzis (OV) tiek ierindots astotajā vietā pēc saslimstības un mirstības rādītājiem sievietēm, izraisot 5 % ar vēzi saistīto nāves gadījumu (Sung et al., 2021). Latvijā KV un OV veido ievērojamu veselības aprūpes slogu, ik gadu nosakot aptuveni 1200 jaunu KV un 300 OV diagnožu (*CDPC*, 2020).

Ir aprēķināts, ka aptuveni 5–10 % no visiem KV un 10–15 % no visiem OV gadījumiem ir pārmantoti, jo tie tiek saistīti ar patogēniem variantiem (PV) vēža predispozīcijas gēnos, it īpaši *BRCA1* un *BRCA2* gēnos (Angeli et al., 2020; Leitsalu et al., 2021). Pārmantojami PV *BRCA1* gēnā ir atzīti par visizplatītāko ģenētisko predispozīciju gan KV, gan OV. Pētījumu dati liecina, ka KV un OV attīstības risks līdz 80 gadu vecumam sievietēm ar *BRCA1* PV variē no 60 % līdz 75 % KV gadījumā un no 34 % līdz 44 % OV gadījumā (Barnes et al., 2020; Borde et al., 2022; Rebbeck et al., 2015). Šie atšķirīgie dati liecina par nepilnīgu penetranci, kas nozīmē, ka tikai daļai *BRCA1* PV nesēju dzīves laikā attīstīsies KV vai OV, padarot ģenētisko konsultēšanu un riska novērtēšanu izaicinošu. Penetrance apzīmē varbūtību, ka indivīds, kuram ir specifiski ģenētiski PV, attīstīs konkrētu pazīmi vai slimību, šajā gadījumā KV vai OV. Līdz ar to tiek uzskatīts, ka šo fenomenu ietekmē citi ģenētiskie faktori (Chen et al., 2020; Downs et al., 2019; Narod, 2002).

Pašlaik indivīda risku saslimt ar KV vai OV izvērtē, balstoties uz personīgo anamnēzi vai pirmās pakāpes radnieku esamību ar konkrētu vēža diagnozi, kā arī ar vecumu saistītiem kritērijiem, kam seko skrīnings, lai identificētu pārmantojamus *BRCA1* PV ar ciltstēva efektu (Jürgens et al., 2022). Tomēr, ņemot vērā *BRCA1* PV nepilnīgo penetranci, novērtējumā būtu jāiekļauj

arī citi penetranci modificējoši faktori. Tā kā profilaktiskās procedūras ir invazīvas un var radīt smagas psiholoģiskas un fizioloģiskas sekas, ģenētiskajā konsultēšanā *BRCA1* PV nesējās ir ļoti svarīgi precīzi novērtēt konkrētā indivīda vecumam atbilstošu vēža attīstības risku. Uzlabota riska prognozēšana var palīdzēt identificēt augsta riska sievietes, kurām varētu būt noderīga agrīna klīniska iejaukšanās, un zema riska sievietes, kuras varētu izlemēt atlikt profilaktiskās procedūras vai profilaksi ar medikamentu palīdzību (Borde et al., 2022; Kuchenbaecker, McGuffog et al., 2017).

Tāpēc šīs disertācijas mērķis bija sniegt ieguldījumu pētījumos par potenciālajiem ģenētiskajiem KV vai OV riska modifikatoriem *BRCA1* PV nesējās, īpaši pievēršoties reģionam specifiskajiem *BRCA1* PV Latvijas populācijā (c.4035del un c.5266dup). Mērķa sasniegšanai tika izmantota uz hipotēzi balstīta mērķtiecīga kandidātģēnu pieeja, koncentrējoties uz *BRCA1* un *CHEK2* dubultheterozigotu analīzi, kam sekoja uz datiem balstīta genoma mēroga asociāciju analīze (*GWAS*). Papildus tika pētīta un salīdzināta divu nesēn izstrādātu genoma mēroga poligēno riska modeļu (*PRS*) (*BayesW* vs. *BayesRR-RC*) efektivitāte, lai novērtētu kopējo ģenētisko risku sievietēm, kurām ir identificēts viens no šiem diviem visbiežāk sastopamajiem pārmantotajiem *BRCA1* PV. Šīs disertācijas mērķis bija novērtēt KV vai OV attīstības risku, ko ietekmē citi ģenētiskie faktori.

Darba mērķis

Identificēt ģenētiskos faktoros, kas varētu ietekmēt divu Latvijā izplatītāko *BRCA1* gēna patogēno variantu (c.4035del un c.5266dup) penetranci pētījuma kohortā.

Darba uzdevumi

Promocijas darba mērķa sasniegšanai tika izvirzīti vairāki uzdevumi:

1. Novērtēt trīs *CHEK2* gēna patogēno/iespējami patogēno variantu ietekmi uz *BRCA1* gēna patogēno variantu (c.4035del un c.5266dup) penetranci pētījuma kohortā.
2. Veikt genoma mēroga asociāciju analīzi (*GWAS*) krūts vēža pacientēm, lai identificētu ģenētiskos variantus, kas ietekmē *BRCA1* gēna patogēno variantu (c.4035del un c.5266dup) penetranci pētījuma kohortā.
3. Veikt genoma mēroga asociāciju analīzi (*GWAS*) olnīcu vēža pacientēm, lai identificētu ģenētiskos variantus, kas ietekmē *BRCA1* gēna patogēno variantu (c.4035del un c.5266dup) penetranci pētījuma kohortā.
4. Novērtēt saistību starp jaunveidotiem genoma mēroga poligēnā riska modeļiem (PRS) un krūts vai olnīcu vēža risku *BRCA1* patogēno variantu (c.4035del un c.5266dup) nesējās pētījuma kohortā.

Darba hipotēze

Reģionam specifisko *BRCA1* gēna patogēno variantu (c.4035del un c.5266dup) penetranci pētījuma kohortā ietekmē citi ģenētiskie faktori.

Darba novitāte

Līdz šim nav izpētīti ģenētiskie faktori, kas ietekmē konkrētu *BRCA1* gēna patogēno variantu nepilnīgu penetranci. Šis ir pirmais ģenētisko modifikatoru pētījums, kas veikts tik lielā kohortā, kura atlasīta pēc Latvijas populācijai un reģionam specifiskiem *BRCA1* patogēnajiem variantiem. Pētījumā tika izmantota ģenētiski homogēna kohorta, kas atlasīta, balstoties uz diviem biežāk sastopamajiem *BRCA1* gēna ciltstēva variantiem (c.4035del un c.5266dup).

1. Materiāli un metodes

1.1. Pētījuma kohorta

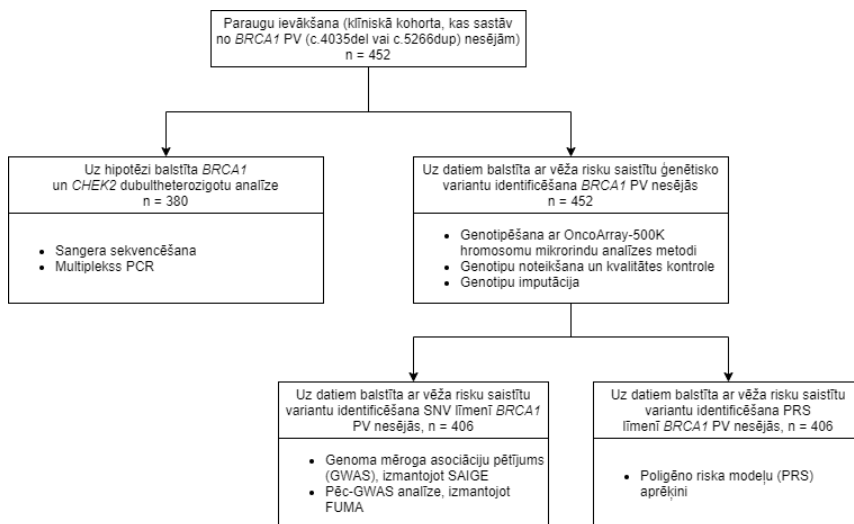
Pētījuma kohorta sastāvēja no 452 sievietēm, kuras tika atlasītas pēc diviem pārmantojamiem *BRCA1* PV – NM_007294.4:c.4035del (rs80357711, iepriekš minēts kā c.4154delA) un NM_007294.4:c.5266dup (rs80357906, iepriekš minēts kā c.5382insC). Pētījuma dalībnieces bija klīniskā kohorta, kas tika nepārtraukti rekrutēta no 2002. līdz 2022. gadam. Dalībnieces bija ≥ 18 gadu vecas, un viņām Paula Stradiņa Klīniskās universitātes slimnīcas Krūts ķirurģijas nodaļā tika veikta ģenētiskā testēšana uz pārmantotu krūts un olnīcu vēža (*HBOC*) sindromu. Dalībnieces tika iedalītas trijās grupās: ar primāri diagnosticētu KV ($n = 196$), ar primāru OV ($n = 129$) vs. grupa bez vēža diagnozes ($n = 127$). Dalībnieču vecums tika cenzēts rekrutēšanas brīdī, un dati par turpmāko novērošanu nebija pieejami. Rekrutēšanas brīdī nevienai no dalībniecēm nebija veikta abpusēja risku samazinoša mastektomija (*BRRM*) vai abpusēja risku samazinoša salpingooforektomija (*BRRSO*). DNS tika izolēta no perifērajām asinīm, izmantojot *FlexiGene* DNA kitu (*Qiagen*, Vācija) saskaņā ar ražotāja protokolu.

Abi testētie varianti ir rāmju nobīdes varianti, kas izraisa priekšlaicīgu stopkodonu, kā rezultātā veidojas saīsināts (c.5266dup) vai samazinātā daudzumā (c.4035del) *BRCA1* proteīns. Abi varianti ir klasificēti kā patogēni, pamatojoties uz Amerikas Medicīniskās ģenētikas un genomikas koledžas (*ACMG*) kritērijiem (Richards et al., 2015), un to bioloģiskā ietekme ir proteīna funkcijas zudums (*LoF*).

Pētījuma datu analīzes darba plūsma ir parādīta 1.1. attēlā.

Lai veiktu šo pētījumu, tika saņemts atzinums no Latvijas Centrālās medicīnas ētikas komitejas par pētījuma atbilstību ētikas prasībām (protokola Nr. 2/18-09-19 un papildinājuma Nr. 01-29.1.2/282). Papildus tika saņemts

apstiprinājums no Genoma izpētes padomes (protokola Nr. A-1/18-10-19). No Igaunijas Bioētikas un cilvēku pētījumu komitejas (Igaunijas Sociālo lietu ministrijas) tika saņemts apstiprinājums par Igaunijas references datu izmantošanu šajā pētījumā (apstiprinājuma Nr. 1.1-12/624 kopā ar papildinājumu Nr. 1.1-12/1478).



1.1. attēls. Promocijas darbā veiktās analīzes darba plūsmas shēma

Katra pētījumā iesaistītā dalībiece bija sniegusi rakstisku informētu piekrišanu par klīniskās un ģenētiskās informācijas izmantošanu pētījuma ietvaros.

1.2. *BRCA1* un *CHEK2* dubultheterozigotu analīze

Uzsākot pētījumu 2019. gadā, tika veikta uz hipotēzi balstīta *BRCA1* un *CHEK2* dubultheterozigotu analīze 380 dalībniecēm, kuru paraugi bija atlasīti līdz pētījuma sākumam (skatīt 1.1. attēlu). *CHEK2* gēna patogēnie/ticami patogēnie un riska varianti (Pavlovica et al., 2022) (splaisa saita variants

NM_007194.4:c.444+1G>A, p.(?), rs121908698 un missensa variants NM_007194.4:c.470 T>C, p.(Ile157Thr), rs17879961) tika identificēti ar Sangera sekvenēšanu, izmantojot *BigDye® Terminator v3.1 Cycle Sequencing* kitu (*Applied Biosystems*, ASV) un praimerus, kā iepriekš aprakstīts (Cybulski et al., 2004). Sekvencēšanas rezultāti tika analizēti, izmantojot ģenētisko analizatoru 3500 (*Applied Biosystems*, ASV). Datu apstrāde un rediģēšana tika veikta ar sekvencēšanas analīzes (*Sequencing Analysis Software*) un *SeqScape™* programmatūru (*Applied Biosystems*, ASV), izmantojot genoma references konsorcijs cilvēka genomu 37 (*GRCh37*)/hg19 (publicēts 2009. gadā). Lai gan pašlaik ir pieejams arī jaunāks references genoms (*GRCh38*), pētījumā tika izmantota vecākā versija, lai atvieglotu rezultātu salīdzināšanu ar iepriekšējiem pētījumiem, kā arī lai nodrošinātu analīzes rīku saderību un samazinātu iespējamās kļūdas, kas saistītas ar pāreju starp genoma versijām. Lai noteiktu PV NM_007194.4:c.(908+1_909-1)_(1095+1_1096-1)del *CHEK2* gēnā, kas izraisa 9.–10. eksonu delēciju (literatūrā minētu arī kā del5395), tika izmantota multipleksas polimerāzes ķēdes reakcijas (PCR) metode (*Veriti, Applied Biosystems*, ASV), kā aprakstīts iepriekšējos rakstos (Cybulski et al., 2007; Plonis et al., 2015). PCR reakcijas produkti tika atdalīti, izmantojot 2 % agarozes gela elektroforēzi. Atklātās delēcijas multipleksas PCR pozitīvos paraugos tika apstiprinātas, izmantojot Sangera sekvencēšanas metodi. Detalizēta informācija par šo metodi ir aprakstīta iepriekš (Cybulski et al., 2006).

1.3. Paraugu genotipēšana ar *OncoArray-500K* hromosomu mikrorindu analīzes metodi

Pētījumā atlasītie 452 paraugi tika genotipēti Rīgas Stradiņa universitātes Onkoloģijas un molekulārās ģenētikas institūtā no 2019. līdz 2022. gadam, izmantojot *Infinium OncoArray-500K* hromosomu mikrorindu analīzes metodi (Illumina, Sandjago, CA, ASV). Hromosomālās mikrorindu analīzes sistēma ietver aptuveni 500 000 marķieru jeb viena nukleotīda variantu (*SNV*), no kuriem

aptuveni 250 000 ir bieži *SNV*, kas sastāda plašu genoma mēroga bāzi. Pārējie marķieri ir ģenētiskie varianti, kas saistīti ar KV, OV un citiem vēža veidiem, un tie galvenokārt ir atklāti, izmantojot iepriekšējās *GWAS* un citas metodes (Guo et al., 2015; Michailidou et al., 2015; Michailidou et al., 2013). Konkrētā hromosomu mikrorindu analīzes sistēma ir izstrādāta, sadarbojoties vadošajiem ekspertiem no *OncoArray* konsorcijs, tostarp krūts vēža asociācijas konsorcijs (*BCAC*), *BRCA1/2* modificējošo faktoru pētnieku konsorcijs (*CIMBA*) un olnīcu vēža asociācijas konsorcijs (*OCAC*).

1.4. Genotipu noteikšana un kvalitātes kontrole (QC)

Iegūto datu analīzē tika izmantots modificēts genotipēto paraugu kvalitātes kontroles (*QC*) process, kas detalizēti aprakstīts citur (Guo et al., 2014). Būtībā tas ietvēra uz paraugiem un variantiem balstītus *QC* soļus, kas tika veikti, galvenokārt izmantojot *GenomeStudio* programmatūru (Illumina, genotipēšanas modulis v2.0.5) un komandrindās balstītu programmu *PLINK* v1.07 un v1.9 (Purcell et al., 2007).

Pirmkārt, parauga genotipa noteikšana tika veikta, izmantojot *GenomeStudio* programmatūru, kurā notika automātiska variantu sagrupēšana. Tālāk manuāli tika izslēgti indivīdi ar genotipēšanas kvalitātes rādītāju zem 98 % vai neatbilstošu dzimumu. Savukārt varianti, kuru *GenTrain* rādītājs bija zemāks par 0,7, tika pārbaudīti un, ja iespējams, manuāli atkārtoti grupēti pirms galīgās eksportēšanas *PLINK* formātā.

Pēc eksportēšanas variantu pozīcijas tika atjauninātas atbilstoši cilvēka references genomam *GRCh37/hg19* un visi varianti tika mainīti no TOP pavediena uz hg19 pozitīvo pavedienu, izmantojot *GSAMD-24v1-0_20011747_A1-b37.strand.RefAlt.zip* failus, kas pieejami vietnē <https://www.well.ox.ac.uk/~wrayner/strand/>.

Lai nodrošinātu datu kvalitāti, tika veikti vairāki *QC* soļi: dzimumu neatbilstība tika pārbaudīta, izmantojot *PLINK* komandu *check-sex*, kas aprēķina X hromosomas inbrīdīga rādītāju, un paraugkopā tika saglabātas sievietes ar inbrīdīga rādītāju $< 0,2$. Rasu neatbilstība tika novērtēta, izmantojot principālo komponentu (PC) analīzi ar programmu *EIGENSOFT* (Price et al., 2006) un 687 informatīvos priekšteču marķierus (*AIMs*), izmantojot definēto sliekšni, kas augstāks par vidējo rādītāju ± 6 standartnovirzes (SD). Radniecība un iespējamie dublikāti tika noteikti, izmantojot pāru identitātes pēc izcelsmes (*IBD*) aprēķinus datos, kuros tika izslēgti marķieri, kas atrodas nelīdzsvarotā saistībā (*LD*), lai uzlabotu analizēto marķieru neatkarību. Tika izslēgti iespējamie dublikāti ar *PI_HAT* vērtību tuvu 1. Novirzes no Hārdija–Veinberga likuma (*HWE*) tika identificētas, izmantojot p vērtības sliekšni $< 1 \times 10^{-7}$ sievietēm bez audzēja diagnozes un sliekšni $< 1 \times 10^{-12}$ sievietēm ar KV vai OV diagnozi. Šajā solī tika izslēgti 503 *SNV*. Papildus tika izslēgti arī paraugi ar ekstrēmu heterozigotitāti ($\pm 4,89$ SD no vidējās vērtības) un inbrīdīga koeficientu $> 0,1$, atstājot 406 paraugus tālākām analīzēm.

1.5. Genotipu imputācija

Lai veiktu imputāciju, tika izslēgti papildus *SNV* ar retās alēles biežumu (*MAF*) $< 0,01$. Trūkstošie genotipi tika imputēti, izmantojot Igaunijas populācijas dziļa pārklājuma visa genoma sekvencēšanas (*WGS*) datus ($n = 2244$) kā references paneli, kā aprakstīts iepriekš (Mitt et al., 2017). Imputācijai tika īstenota divpakāpju pieeja: fāzēšana ar *EAGLE* (Loh et al., 2016) un imputācija ar *BEAGLE* (Browning et al., 2018). Rezultātā tika ģenerēti genotipi aptuveni 38 miljoniem *SNV*. Pēc imputācijas tika veikta *QC*, izslēdzot retus *SNV* ar *MAF* $< 0,01$ un devas R-kvadrātā (*DR2*) rādītāju $< 0,8$. Filtrētie dati saturēja 7 911 505 labas kvalitātes *SNV* turpmākām analīzēm.

1.6. Genoma mēroga asociāciju analīze (GWAS), izmantojot SAIGE

Pēc datu QC filtrēšanas un imputācijas soļiem asociāciju analīzei bija pieejami 406 indivīdi. Asociāciju analīze tika veikta, izmantojot programmatūras R v4.0.2 (R. C. Team, 2020) pakotni SAIGE v0.38 (Chen et al., 2016), lai pielietotu jauktu loģistiskās regresijas modeli. Modelis tika koriģēts atbilstoši radniecībai, pirmajiem 4 PC, vecumam rekrutēšanas/slimības sākuma brīdī un *BRCA1* PV. Šajā pētījumā radniecība tika pielāgota, lai mazinātu viltus pozitīvu asociāciju risku un nodrošinātu, ka pārbaudītie ģenētiskie varianti patiesi ir saistīti ar pētāmo iznākumu (t.i., KV vai OV), nevis tiek ietekmēti no radnieciskām attiecībām. Jaukta loģistiskās regresijas modeļa ieviešana kopā ar radniecības un citu kovariātu pielāgošanu palīdz kontrolēt iespējamus neobjektivitātes avotus, nodrošinot ticamākus rezultātus. Asociācijas analīzei tika izmantots genoma mēroga nozīmīguma sliekšnis $p < 5 \times 10^{-8}$, kas pēc-GWAS analīzēs tika samazināts līdz iespējamam nozīmīguma sliekšnim $p < 1 \times 10^{-6}$.

1.7. Pēc-GWAS analīze, izmantojot brīvpiekļuves platformas FUMA un VEP

Funkcionālās kartēšanas un anotācijas (FUMA) platforma tika izmantota, lai anotētu, prioritizētu, vizualizētu un interpretētu GWAS rezultātus. Lai identificētu neatkarīgus nozīmīgus SNV, no GWAS rezultātiem tika atlasīti SNV, kuru p vērtība bija mazāka vai vienāda ar 1×10^{-6} un $r^2 < 0,6$. Turklāt, lai izšķirtu vadošos SNV no neatkarīgiem nozīmīgiem SNV, tika izmantots pāra salīdzinājuma SNV sliekšnis $r^2 < 0,1$. Tālāk tika noteikti genomiskie riska lokusi, kuros esošie SNV bija augstā LD ar neatkarīgajiem nozīmīgajiem SNV un to r^2 koeficients pārsniedza 0,6. LD bloku apvienošanai vienā genomiskajā riska lokusā tika izmantots maksimālais attālums 250 kb starp LD blokiem. Lai veiktu

šo analīzi, kā references dati tika izmantoti ģenētiskie dati no 1000 genomu projekta 3. fāzes.

Papildus, lai novērtētu *GWAS* identificēto vadošo *SNV* ietekmi uz gēniem, transkriptu un regulatīvajiem reģioniem, tika izmantots *Ensembl* varianta efekta prognozēšanas (*VEP*) rīks (<https://www.ensembl.org/info/docs/tools/vep/index.html>).

SNP2GENE funkcija tika izmantota, lai aprēķinātu *LD* struktūru, raksturotu genomiskos riska lokusus, anotētu *SNV* funkcijas un prioritizētu kandidātgēnus. Pamatojoties uz *SNV* pozicionālo kartēšanu, tika noteikti tuvākie kandidātgēni, kas atradās ne vairāk kā 10 kb attālumā.

Papildus tika veikta ekspresijas kvantitatīvo pazīmju lokusu (*eQTL*) kartēšana, lai noteiktu saistību starp *GWAS* identificētajiem *SNV* un izmaiņām gēnu ekspresijā. Šī analīze palīdz izprast identificēto ģenētisko variantu funkcionālās sekas un sniegt ieskatu iespējamajos bioloģiskajos mehānismos, kas ir novēroto ģenētisko asociāciju pamatā. *eQTL* kartēšanai tika izmantoti dati no diviem audu tipiem, t.i., genotipa–audu ekspresijas (*GTE_x*) projekta v8 krūts un *GTE_x* v8 olnīcu audiem. Par nozīmīgām tika uzskatītas tikai tās *eQTL* vērtības, kuru viltus atklāšanas kļūda (*FDR*) bija mazāka par 0,05, un tās tika izmantotas *SNV* kartēšanai gēnos.

1.8. Poligēnā riska modeļa (PRS) aprēķini

Šajā pētījumā izmantotajos PRS aprēķinos tika iekļauta informācija par 2 174 072 *SNV*, kas ir sastopami gan Apvienotās Karalistes Biobankas (<https://www.ukbiobank.ac.uk/> (Bycroft et al., 2018)), gan Igaunijas Biobankas indivīdos (<https://genomics.ut.ee/en/content/estonian-biobank> (Mitt et al., 2017)). Šie PRS tika izstrādāti, izmantojot datus no 428 747 Apvienotās Karalistes Biobankas indivīdiem un 105 000 Igaunijas Genoma centra dalībniekiem (Orliac et al., 2022). Šajā promocijas darbā veiktajos aprēķinos tika

iekļauta informācija par 2 041 044 *SNV*, jo datu kopā iztrūka informācija par 133 028 *SNV*. Visiem PRS aprēķiniem tika izmantota *PLINK* v2.00 funkcija *score*.

1.9. Statistiskā analīze

Statistiskajai analīzei tika izmantotas programmatūras R v4.0.2 (*R Core Team*, Vīne, Austrija) (R. C. Team, 2020) un *RStudio* v1.3.1093 (*RStudio Team*, Bostona, MA, ASV) (R. Team, 2020). Visi veiktie statistiskie testi bija divpusēji, un par statistiski nozīmīgām tika uzskatītas p vērtības zem 0,05.

Lai sasniegtu konkrētus pētījuma mērķus, tika izmantotas dažādas statistikas metodes un R pakotnes. Piemēram, lai novērtētu vecuma sadalījuma atšķirības starp pētījuma grupām, tika izmantots Kruskala–Valisa tests (R bāzes pakotne *stats*), kam sekoja *posthoc* pāru salīdzinājumi ar Vilkoksona rangū summas testu ar Bonferroni korekciju. *BRCA1* un *CHEK2* dubultheterozigotu izplatība un to saistība ar KV un/vai OV tika novērtēta, izmantojot Fišera precīzo testu, lai noteiktu izredžu attiecību (OR) un to statistisko nozīmīgumu. Turklāt tika izmantota *Bioconductor* pakotne *Survival*, 3.2-3 versija (Therneau, 2020), lai izpētītu PV ietekmi uz kumulatīvo KV un/vai OV risku, izmantojot Kaplana–Meiera aplēses, savukārt līknes atšķirības tika novērtētas, izmantojot *Log-rank* testu. Lai prognozētu kumulatīvo risku (varbūtību līdz notikumam), tika veikta Koksas regresijas analīze.

Saistība starp PRS un KV un/vai OV attīstības risku *BRCA1* PV nesējās tika novērtēta, izmantojot binomiālās loģistikās regresijas modeli. Iznākuma mainīgajam bija trīs kategorijas: 0 (nav vēzis), 1 (KV) un/vai 2 (OV). Modelis tika pielāgots vecumam, vecuma kvadrātam, *BRCA1* PV (c.4035del vai c.5266dup) un pirmajiem diviem PC. OR un to 95 % ticamības intervāls (TI) tika aprēķināti, izmantojot R pakotni *Epi* (Carstensen, 2022). Lai izvēlētos optimālo binomiālās loģistikās regresijas analīzes modeli, tika veikta uztvērēja darbības raksturlienes (ROC) analīze, izmantojot R pakotni *pROC* (Robin et al., 2011).

1.10. Datu pieejamība

GWAS analīzes kopsavilkuma dati (angl. *summary statistics*) būs pieejami <https://dataverse.rsu.lv/> repozitorijā.

2. Rezultāti

2.1. Pētījuma kohortas raksturojums

2.1.1. Pacientu raksturojums

Pētījuma kohorta sastāvēja no 452 sievietēm, kuras bija *BRCA1* PV (c.4035del vai c.5266dup) nesējas. Šīm sievietēm bija diagnosticēts KV, OV, vai tās bija bez vēža diagnozes rekrutēšanas brīdī. Pētījuma kohorta sastāvēja no 196 sievietēm (43,4 %) ar diagnosticētu KV, 129 sievietēm (28,5 %) ar diagnosticētu OV un 127 sievietēm (28,1 %) bez vēža diagnozes, kas kalpoja kā grupa bez vēža diagnozes salīdzinājumam. Vidējais vecums KV vai OV diagnosticēšanas brīdī bija attiecīgi 46,52 gadi (diapazons 25–92 gadi, SD = 11,71) un 50,62 gadi (diapazons 27–79 gadi, SD = 8,80). Ar vēzi neskartās grupas vidējais vecums bija 38,36 gadi (diapazons 18–73 gadi, SD = 11,05). Tika veikti pacientu vecuma salīdzinājumi pa pāriem starp dažādām grupām, izmantojot Vilksona rangu summas testu ar nepārtrauktības korekciju, un tie atklāja statistiski nozīmīgu $p < 0,01$ visās trijās grupās, norādot uz būtiskām vecuma atšķirībām starp katru grupu. Lai korekti veiktu turpmākās analīzes, šīs vecuma atšķirības starp KV un OV grupām un grupu bez vēža diagnozes tika pielāgotas un standartizētas. Pētījuma kohortas raksturojums ir apkopots 2.1. tabulā, un tas ir būtisks, lai izprastu kohortas daudzveidību un izdarītu secinājumus par *BRCA1* PV ietekmi uz vēža attīstību.

Pēc daudzpakāpju *QC* soļiem un plašas datu filtrēšanas pētījuma datu kopa tika samazināta līdz 406 paraugiem. Galīgā pētījuma kohorta, kas tika izmantota turpmākajai *GWAS* un PRS analīzei, sastāvēja no 171 sievietes (42,1 %) ar KV diagnozi, 121 sievietes (29,8 %) ar OV diagnozi un 114 sievietēm (28,1 %) bez vēža diagnozes. Vidējais slimības sākuma vecums bija 46,67 gadi (diapazons 25–92 gadi) KV gadījumā un 50,55 gadi (diapazons 27–79 gadi) OV gadījumā.

Pētījuma kohortas raksturojums

	Kopā	BRCA1:c.4035del	BRCA1:c.5266dup
Pētījuma kohorta	452	173 (38,28 %)	279 (61,72 %)
Krūts vēzis	196 (43,36 %)	53 (11,73 %)	143 (31,64 %)
Olnīcu vēzis	129 (28,54 %)	69 (15,27 %)	60 (13,27 %)
Nav vēzis	127 (28,10 %)	51 (11,28 %)	76 (16,81 %)
Vidējais vecums	45,40 ± 11,72	47,67 ± 12,02	43,99 ± 11,35
Krūts vēzis*	46,52 ± 11,71	49,68 ± 12,56	45,34 ± 11,19
Olnīcu vēzis*	50,62 ± 8,80	52,00 ± 9,57	49,03 ± 7,58
Nav vēzis*	38,36 ± 11,05	39,73 ± 10,64	37,45 ± 11,30

* Apzīmē statistiski nozīmīgu ($p < 0,01$) vecuma atšķirību starp visām 3 pētījuma grupām.

2.1.2. BRCA1 PV c.4035del un c.5266dup penetrance pētījuma kohortā

Pētījuma kohorta tika sadalīta divās apakšgrupās, balstoties uz diviem specifiskiem *BRCA1* gēna PV (c.4035del un c.5266dup) ar ciltstēva efektu. Kopējā pētījuma kohorta sastāvēja no 173 sievietēm, kuras bija *BRCA1*:c.4035del PV nesējas (53 sievietes KV grupā, 69 sievietes OV grupā un 51 sieviete grupā bez vēža diagnozes), un 279 sievietēm, kuras bija *BRCA1*:c.5266dup PV nesējas (143 sievietes KV grupā, 60 sievietes OV grupā un 76 sievietes grupā bez vēža diagnozes), kā parādīts 2.1. tabulā.

Tālāk tika aprēķināta *BRCA1* PV penetrance, kas tiek definēta kā indivīdu proporcija, kuri ir ar slimību saistīto PV nesēji un kuriem attīstās atbilstošais KV vai OV fenotips (Cooper et al., 2013). Rezultāti ir sniegti 2.2. tabulā. Pētījuma kohortā aprēķinātā penetrance starp *BRCA1* PV (c.4035del un c.5266dup) nesējām bija attiecīgi 31 % KV un 40 % OV gadījumā un 51 % KV un 22 % OV gadījumā.

BRCA1 PV c.4035del un c.5266dup penetrance pētījuma kohortā

BRCA1 PV	Krūts vēzis (%)	Olnīcu vēzis (%)
c.4035del	30,64	39,88
c.5266dup	51,25	21,51

PV – patogēns variants.

2.1.3. Ar vecumu saistītā KV vai OV kumulatīvā incidence BRCA1 c.4035del un c.5266dup PV nesējās

Tālāk tika veikta Koksas proporcionālo risku regresijas analīze, lai noteiktu saistību starp *BRCA1* PV (c.4035del vai c.5266dup) un laiku līdz notikumam (vēža diagnozei) šajā pētījuma kohortā, kura sastāvēja no 452 indivīdiem. No tiem 325 indivīdiem bija diagnosticēts vēzis (196 KV gadījumi un 129 OV gadījumi).

Analīze atklāja statistiski nozīmīgu saistību starp *BRCA1*:c.5266dup PV un vēža rašanās vecumu ar regresijas koeficientu 0,3626 ($p = 0,00169^{**}$). Aprēķinātā riska attiecība (HR) *BRCA1*:c.5266dup PV bija 1,437 (95 % TI: 1,15–1,80), kas liecina, ka indivīdiem ar šo variantu ir par 43,70 % lielāks risks saslimt ar KV vai OV jaunākā vecumā salīdzinājumā ar personām ar otru *BRCA1*:c.4035del PV.

Konkordances indekss, kas ir izmantotā modeļa prognozēšanas precizitātes rādītājs, bija 0,562 (standartklūda (SE) = 0,015), kas norāda uz vidēji labu prognozēšanas spēju. Papildu statistiskie testi konsekventi apstiprināja šīs saistības nozīmīgumu. Veicot ticamības koeficienta (angl. *likelihood ratio*), *Wald* un *Log-rank* testus, tika konstatēta nozīmīga saistība starp *BRCA1*:c.5266dup PV un vēža attīstību (attiecīgi $p = 0,001$, $p = 0,002$ un $p = 0,002$). Šie rezultāti liecina, ka *BRCA1*:c.5266dup PV ir statistiski nozīmīgs agrīnākas vēža attīstības prognozētājs.

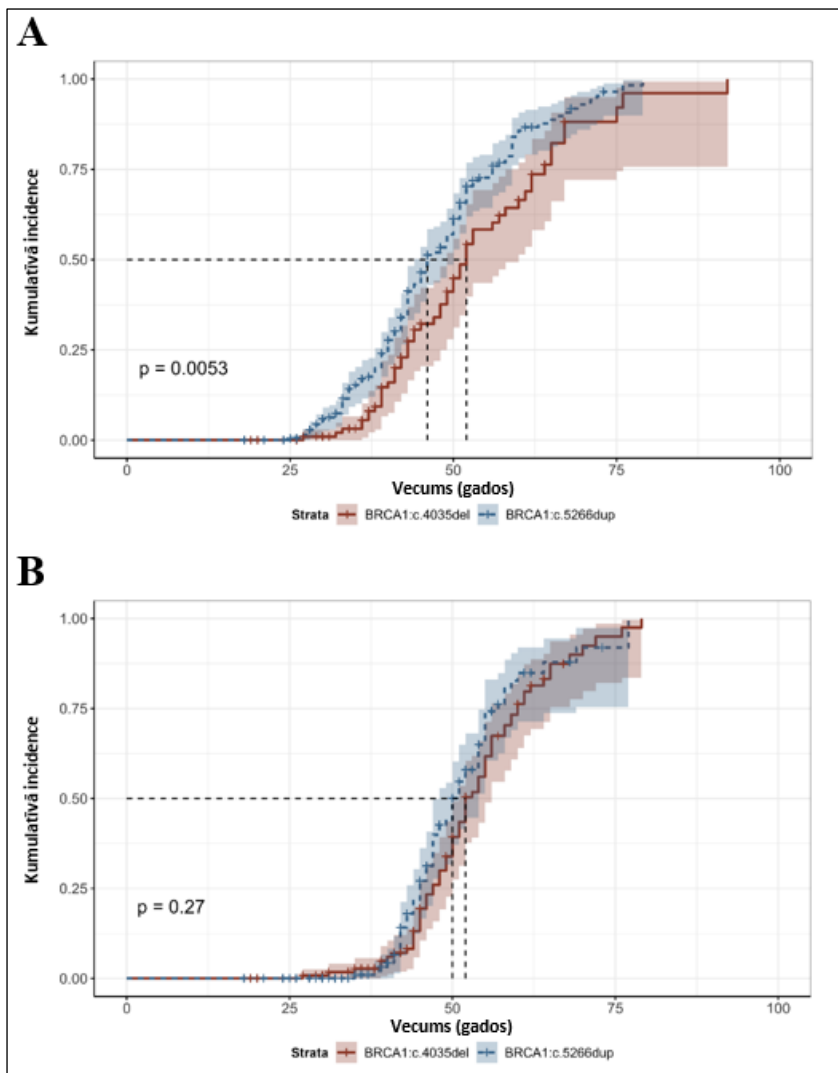
Turpmākajā analizē tika pētīta *BRCA1:c.5266dup PV* ietekme uz *KV* un *OV* grupām atsevišķi (skatīt 2.1. attēlu).

Katrai grupai atsevišķi tika piemēroti Koksas proporcionālo risku regresijas modeļi un iegūti šādi rezultāti: *KV* grupas analizē, kura sastāvēja no 323 indivīdiem ar 196 vēža gadījumiem, Koksas regresijas analīze atklāja nozīmīgu saistību starp *BRCA1:c.5266dup PV* un *KV* rašanās vecumu.

Aprēķinātā *BRCA1:c.5266dup PV HR* bija 1,564 (95 % *TI*: 1,14–2,15), kas norāda uz 56,40 % augstāku risku saslimt ar *KV*, salīdzinot ar sievietēm, kuras ir otra *BRCA1 PV* (c.4035del) nesējas. Papildus veiktie statistiskie testi arī apstiprināja šīs saistības nozīmīgumu. Veicot ticamības koeficienta testu, tika iegūta $p = 0,005$, *Wald* un *Log-rank* testu rezultāts bija $p = 0,006$, un konkordances indekss bija 0,56, norādot uz vidēju modeļa prognozēšanas spēju.

Savukārt *OV* grupas analizē, kurā tika iekļauti 256 indivīdi ar 129 vēža gadījumiem, Koksas regresijas analīze neuzrādīja līdzīgu tendenci. Saistība starp *BRCA1:c.5266dup PV* un agrāku vēža rašanās vecumu šajā grupā nebija statistiski nozīmīga. Aprēķinātā *HR BRCA1:c.5266dup PV* bija 1,2198 (95 % *TI*: 0,86–1,73) ar $p = 0,265$.

Papildus veicot ticamības koeficienta, *Wald* un *Log-rank* testus, iegūtās *p* vērtības bija aptuveni 0,3, kas norāda, ka pierādījumi nav pietiekami pārliecinoši, lai secinātu, ka *BRCA1:c.5266dup PV* būtiski ietekmē *OV* attīstības vecumu pētījuma kohortā. Šie atšķirīgie rezultāti liecina, ka *BRCA1:c.5266dup PV* ir būtiska nozīme vēža attīstības vecumā *KV* grupā, bet tā ietekme *OV* grupā ir mazāk izteikta.



2.1. attēls. KV un OV kumulatīvā incidence *BRCA1* PV nesējās

Sarkanā līnija attēlo *BRCA1*:c.4035del varianta nesējas; zilā līnija attēlo *BRCA1*:c.5266dup varianta nesējas. **A** – grafiks vizualizē KV attīstības kumulatīvo incidenci *BRCA1* PV nesējās; **B** – grafiks vizualizē OV attīstības kumulatīvo incidenci *BRCA1* PV nesējās. KV – krūts vēzis, OV – olnīcu vēzis.

2.2. Pētījuma dizains: uz hipotēzi balstīta vs. uz datiem balstīta analīze

Šī nodaļa sniedz pārskatu par pētījuma dizaina struktūru, iedalot to divās pieejās – uz hipotēzi balstīts pētījums un uz datiem balstīts pētījums. Šīs divas pieejas būtiski atšķiras ar to, kā tās nosaka un sasniedz pētījuma mērķus. Uz hipotēzi balstītam pētījumam ir raksturīga konkrētas pētījuma hipotēzes izvirzīšana pirms datu vākšanas un analīzes. Turpretī uz datiem balstītam pētījumam ir raksturīga datu izpēte bez iepriekš izvirzītas hipotēzes.

Šajā promocijas darbā tika izmantota hibrīda metodoloģija, kas apvienoja gan uz hipotēzi balstītu, gan uz datiem balstītu pieeju. Šī kombinācija ļāva veikt konkrētas hipotēzes testēšanu, vienlaikus pētot arī jaunas un neparedzētas asociācijas analizētajos datos.

2.2.1. Uz hipotēzi balstīta *BRCA1* un *CHEK2* dubultheterozigotu analīze

Šajā pētījuma sadaļā galvenā uzmanība tika vērsta uz hipotēzi balstītu pieeju, konkrēti pētot individuus, kuros ir novērojama dubultheterozigotitāte *BRCA1* un *CHEK2* gēnos, jo abi minētie gēni ir iesaistīti vienā DNS reparācijas signālceļā. Šī analīze tika veikta pētījuma sākumā, izmantojot datus no 380 indivīdiem.

Analizētie *CHEK2* varianti tika atklāti 13 dubultheterozigotas gadījumos (tostarp c.444+1G>A, n = 1, c.470T>C, n = 11, del5395, n = 1), kā norādīts 2.3. tabulā.

CHEK2 variantu frekvences pētījuma kohortā

Variants un gadījumi	Gadījumu skaits/kopā	Frekvence (%)
c.444+1G>A		
Nav vēzis	1/87	1,15
KV gadījumi	0/132	0,00
OV gadījumi	0/111	0,00
c.470T>C		
Nav vēzis	2/87	2,30
KV gadījumi	3/132	2,27
OV gadījumi	6/111	5,41
del5395		
Nav vēzis	0/87	0,00
KV gadījumi	1/129	0,78
OV gadījumi	0/109	0,00

KV – krūts vēzis; OV – olnīcu vēzis.

Nevienā no analizētajiem paraugiem vienlaikus netika atklāts vairāk kā viens *CHEK2* variants. Lai novērtētu *CHEK2* alēlisko variantu penetranci saistībā ar KV vai OV attīstības risku *BRCA1* PV nesējās, tika salīdzināta šo *CHEK2* variantu izplatība KV un OV grupās pret grupu bez vēža diagnozes pētījuma kohortas ietvaros. Lai gan *CHEK2* variantu izplatība OV grupā bija salīdzinoši augsta (5,41 %), OV riska pieaugums nebija statistiski nozīmīgs (OR = 1,56; 95 % TI: 0,32–9,94; p = 0,73). Arī pētīto *CHEK2* variantu izplatība KV grupā būtiski neatšķīrās no grupas bez vēža diagnozes (OR = 0,88; 95 % TI: 0,15–6,15; p = 1).

Dati par *CHEK2* variantu ietekmi uz vēža attīstības vecumu nebija konsekventi. *BRCA1*:c.4035del PV nesēju vidū neviena pētītā *CHEK2* varianta klātbūtne būtiski neietekmēja vidējo vēža attīstības vecumu (p > 0,3 pēc *Log-rank* testa). Savukārt *BRCA1*:c.5266dup PV nesējās, kurās tika identificēts arī kāds no analizētajiem *CHEK2* variantiem, vidējais OV diagnosticēšanas vecums bija ievērojami zemāks – 8,5 gadu atšķirība salīdzinājumā ar *BRCA1*:c.5266dup PV nesējām bez identificēta *CHEK2* varianta. Aprēķinātā HR

bija 3,93 (95 % TI: 0,93–16,65). Lai gan *Log-rank* tests liecināja par statistiski nozīmīgu atšķirību ($p = 0,043$) un tendence liecināja par saistību starp *BRCA1* un *CHEK2* dubultheterozigotām un jaunāku OV attīstības vecumu, alternatīvā Koksas regresijas analīze nerasniedza statistiski nozīmīgu rezultātu (regresijas koeficients: 1,37, $p = 0,064$).

2.2.2. Uz datiem balstīta ar vēža risku saistītu variantu identificēšana viena nukleotīda varianta (SNV) līmenī *BRCA1* PV nesējās

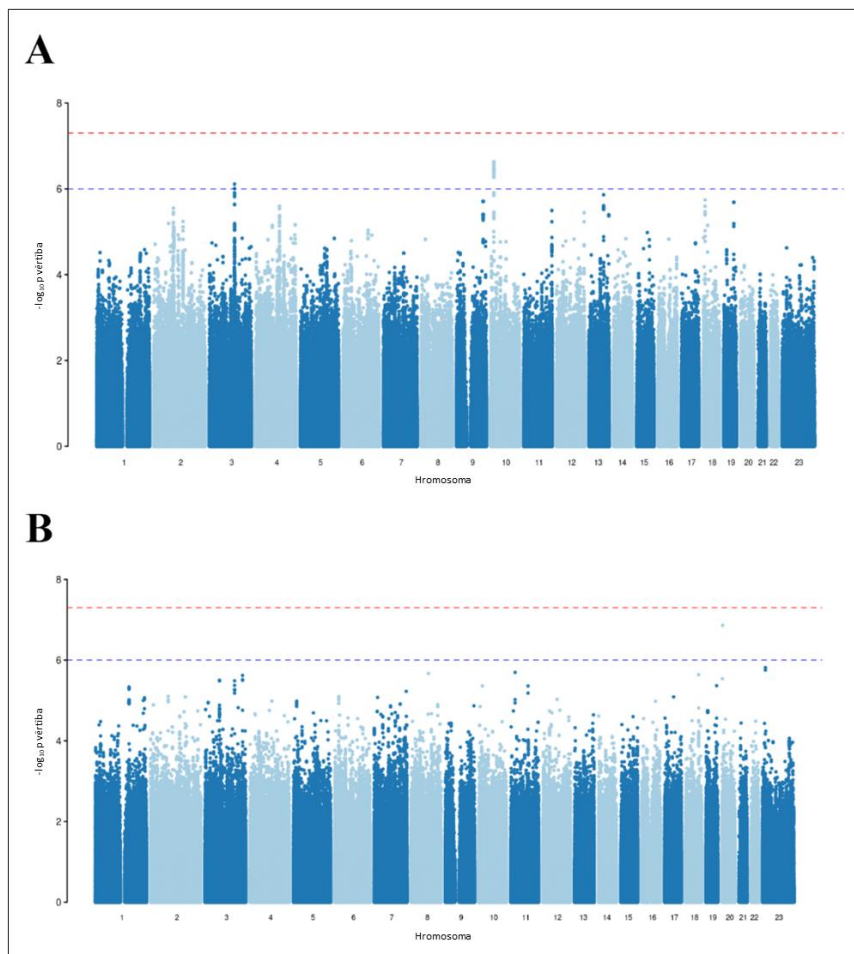
Šajā sadaļā ir pāreja uz datiem balstītu pieeju, kurā galvenā uzmanība tika vērsta uz ģenētisko variantu identificēšanu, kas saistīti ar KV un OV attīstības risku sievietēm, kuras ir *BRCA1* PV nesējās, bez iepriekš izvirzītām hipotēzēm. Pētījumā tika izmantota *GWAS* pieeja, lai identificētu šādus variantus.

Kopumā tika analizēti 7 911 505 *SNV*, lai noteiktu saistību ar KV vai OV attīstības risku 406 *BRCA1* PV nesējās. Pētījuma analītiskajos modeļos tika iekļauti tādi kovariāti kā vecums rekrutēšanas/slimības diagnozes brīdī, radniecība starp dalībniekiem un konkrētā *BRCA1* PV veids.

Būtiskākie rezultāti, kas atspoguļo analizēto *SNV* saistību ar KV vai OV attīstības risku, ir attēloti Manhetenas diagrammās 2.2. attēlā un detalizētāk uzskaitīti šī darba 2.4. tabulā. Atklāto asociāciju noteikšanai tika izmantots stingrs genoma mēroga statistiskā nozīmīguma sliekšnis $p < 5 \times 10^{-8}$, savukārt p vērtības no 5×10^{-8} līdz $\leq 1 \times 10^{-6}$ tika uzskatītas par iespējama nozīmīguma sliekšni. Statistiski nozīmīgākie *SNV*, kas saistīti ar KV attīstības risku, tika identificēti 3. un 10. hromosomā, un no tiem visbūtiskākā saistība bija *SNV*, kas atradās 10. hromosomā (skatīt 2.2. A attēlu). Savukārt nozīmīgākais ar OV saistītais *SNV* bija 20. hromosomā, kā parādīts 2.2. B attēlā.

Lai novērtētu iespējamās novirzes un analizēto datu kvalitāti, tika izveidoti kvantiļu–kvantiļu ($Q-Q$) grafiki, lai novērtētu genomu faktorus KV un OV grupās (šajā kopsavilkumā nav iekļauti). Aprēķinātie inflācijas faktori (λ) bija

0,995 KV un 1,003 OV grupā, kas norāda, ka veiktajā analizē nebija būtiska genomiska inflācija.



2.2. attēls. *SNV* saistība ar KV vai OV attīstības risku

A – Manhetenas diagramma, kurā vizualizētas $-\log_{10}$ vērtības *SNV* asociācijām ar KV attīstības risku. **B** – Manhetenas diagramma, kurā vizualizētas $-\log_{10}$ vērtības *SNV* asociācijām ar OV attīstības risku. Sarkanā līnija apzīmē genoma mēroga nozīmīgumu ($p = 5 \times 10^{-8}$); zilā līnija apzīmē genoma mēroga iespējamo nozīmīgumu ($p = 1 \times 10^{-6}$); 23. hromosoma apzīmē X hromosomu.

Tālāk pēc-GWAS analizē tika izmantota *FUMA* platforma, kurā tika identificēti 18 genomiskie riska lokusi, kas saistīti ar KV, un 21 genomiskais riska lokuss, kas saistīts ar OV attīstības risku. Šajos lokusos atradās 27 neatkarīgi nozīmīgi *SNV* KV grupā un 25 neatkarīgi nozīmīgi *SNV* OV grupā, kas sasniedza iepriekš noteikto iespējama nozīmīguma sliekšni $p < 1 \times 10^{-6}$ un bija savstarpēji neatkarīgi ar $r^2 < 0,6$. Papildus mēs identificējām 1152 kandidāt-*SNV* KV grupā un 633 *SNV* OV grupā, kuri atradās *LD* ar iepriekš minētajiem neatkarīgajiem nozīmīgajiem *SNV*.

2.4. tabulā ir uzskaitīti trīs nozīmīgākie ($p < 1 \times 10^{-7}$) identificētie ģenētiskie varianti, kas izrādīja būtiskāko saistību ar KV vai OV attīstības risku.

2.4. tabula

Visnozīmīgāk saistītie varianti ar KV vai OV attīstības risku

Grupa	rsID	REF	ALT	MAF	P vērtība	Beta	SE	Tuvākais gēns
KV	rs2609813	A	G	0,07952	$2,33 \times 10^{-7}$	-1,26	0,24	<i>FAM107B</i>
KV	rs4688094	G	C	0,4523	$7,76 \times 10^{-7}$	-0,96	0,19	<i>RP11-384F7.1</i>
OV	rs79732499	G	T	0,01789	$1,38 \times 10^{-7}$	-8,09	1,54	<i>C20orf194</i>

KV – krūts vēzis; OV – olnīcu vēzis; rsID – referenes *SNV* ID numurs; REF – referenes alēle; ALT – alternatīvā alēle; MAF – retās alēles biežums; Beta – multivariatīvās lineārās regresijas koeficients; SE – standartkļūda.

Izplatīta prakse pēc-GWAS analīzes ir kandidāt-*SNV* anotācija tuvākajam ģenam. Lēmums ziņot tuvāko gēnu bieži vien ir praktisks, pamatojoties uz pieņēmumu, ka tuvums korelē ar lielāku varbūtību, ka atklātais *SNV* ietekmē gēna funkciju. Tomēr ir būtiski apzināties, ka tuvākais gēns ne vienmēr var būt funkcionālais gēns, kas ietekmē novēroto saistību ar analizēto fenotipu (Watanabe et al., 2017).

Mūsu identificētie vadošie *SNV* ir vērtīgi kandidāti turpmākiem funkcionāliem pētījumiem, nodrošinot pamatu, lai izprastu sarežģītus molekulārus mehānismus, kas ietekmē *BRCA1* PV penetranci.

Statistiski nozīmīgākā saistība ar KV attīstības risku tika novērota rs2609813 variantam ($\beta = -1,26$; $p = 2,33 \times 10^{-7}$; riska alēles G biežums = 0,08). Sīkāka informācija par šo variantu ir pieejama 2.4. tabulā. Šis vadošais variants rs2609813 atrodas 10. hromosomā, un tas ir *FAM107B* (*Family with Sequence Similarity 107 Member B*) proteīnu kodējoša gēna (ENSG00000065809) intronisks variants. Jāatzīmē, ka šajā introniskajā reģionā tika kartēti vēl 56 *SNV*, kas uzrāda augstu *LD* ar vadošo variantu. Pamatojoties uz *VEP* rīku, tiek prognozēts, ka šis introniskais variants ir regulatorā reģiona variants, kas atrodas enhanserī.

Otrs visbūtiskāk saistītais *SNV* ar KV attīstības risku bija rs4688094 ($\beta = -0,96$; $p = 7,76 \times 10^{-7}$; riska alēles C biežums = 0,45), kā parādīts 2.4. tabulā. Atklātais variants atrodas 3. hromosomā iepriekš neaprakstītas garās nekodējošās RNS (*lncRNA*) gēnā *RP11-384F7.1* (ENSG00000243276), un tas uzrāda augstu *LD* ar 295 citiem *SNV*.

Vienīgais variants, kas OV grupā sasniedza iespējama nozīmīguma sliekšni $p < 1 \times 10^{-7}$, bija vadošais variants rs79732499. Šis variants uzrādīja zemāko šajā pētījumā novēroto p vērtību ($\beta = -8,09$; $p = 1,39 \times 10^{-7}$) ar riska alēles T biežumu 0,018 (skatīt 2.4. tabulu). Tas atrodas 20. hromosomas intergēniskā reģionā, un tuvākais kartētais gēns ir *DNAAF9* (*Dynein Axonemal Assembly Factor 9*, agrāk pazīstams kā *C20orf194*), kas ir proteīnu kodējošs gēns (ENSG00000088854). Vadošais variants rs79732499 atrodas *LD* ar četriem citiem *SNV*, kas kartēti šajā gēnā. Pamatojoties uz *VEP* rīku, tiek prognozēts, ka šis intergēniskais variants atrodas starp gēniem regulējošā reģionā (enhanserī).

Tālāk mēs veicām *eQTL* kartēšanu, koncentrējoties uz ģenētisko variantu ietekmi uz gēnu ekspresiju, izmantojot publiski pieejamos *GTEX* krūts un olnīcu

audu datus. *GTEX* datubāze satur 563 genotipētus paraugus, no kuriem tika izmantoti dati no normālu krūts ($n = 396$) un olnīcu ($n = 167$) audu paraugiem. Kartēšana tika veikta, lai atklātu potenciāli funkcionālus variantus mūsu analizētajos datos, prognozētu mērķa gēnus un tos prioritizētu turpmākiem eksperimentāliem pētījumiem. Piemērojot *FDR* sliekšni, kas mazāks par 0,05, un analizējot pētījumā identificētos kandidāt-*SNV*, olnīcu audos netika konstatēti nozīmīgi *SNV* gēnu *cis-eQTL* pāri. Savukārt krūts audos tika novērotas divas nozīmīgas *eQTL* vērtības (skatīt 2.5. tabulu).

2.5. tabula

eQTL rezultāti krūts audos

rsID	REF	ALT	MAF	p vērtība	FDR	NES	Tuvākais gēns
rs10178186	C	T	0,10	$3,83 \times 10^{-7}$	$1,55 \times 10^{-16}$	-0,36	<i>ZNF514</i>
rs434451	T	C	0,035	$2,90 \times 10^{-6}$	0,011	-0,42	<i>SLC1A5</i>

rsID – references *SNV* ID numurs; REF – references alēle; ALT – alternatīvā alēle; *MAF* – retās alēles biežums; *FDR* – viltus atklāšanas kļūda; *NES* – normalizētais efekta lielums, ko definē kā lineārās regresijas slīpumu un aprēķina kā alternatīvās alēles (ALT) ietekmi attiecībā pret references alēles (REF) ietekmi cilvēka references genomā (pieņemot, ka *eQTL* efekta alēle ir ALT alēle).

Visnozīmīgākā saistība identificēta vadošajam *SNV* rs10178186 ar nekoriģētu p vērtību $3,83 \times 10^{-7}$ un riska alēles T frekvenci 0,10 (skatīt 2.5. tabulu). Šis variants kopā ar 99 citiem variantiem, kas atrodas augstā *LD*, tika kartēts proteīnu kodējošā gēnā *ZNF514* (*Zinc Finger Protein 514*) (ENSG00000144026), kas atrodas 2. hromosomā. Normalizētais efekta lielums (*NES*) = -0,36 norāda uz negatīvu saistību starp rs10178186 variantu un *ZNF514* gēna ekspresiju.

Otra statistiski nozīmīgā *eQTL* asociācija tika identificēta vadošajam *SNV* rs434451 ar nekoriģētu p vērtību $2,90 \times 10^{-6}$ un riska alēles C biežumu 0,96 (skatīt 2.5. tabulu). Interesanti, ka šis bija vienīgais variants, kas tika kartēts proteīnu kodējošā gēnā *SLC1A5* (*Solute Carrier Family 1 Member 5*)

(ENSG00000105281), kas atrodas 19. hromosomā. $NES = -0,42$ norāda uz negatīvu saistību starp rs434451 variantu un *SLCIA5* gēna ekspresiju.

2.2.3. Uz datiem balstīta ar vēža risku saistītu variantu identificēšana poligēnu riska variantu (PRS) līmenī *BRCA1* PV nesējās

Šajā pētījumā tika izmantoti četri dažādi PRS apvienotie modeļi, kas apzīmēti ar modelis1 līdz modelis4, lai novērtētu ģenētisko risku saslimt ar KV vai OV *BRCA1* PV nesējās. Jāatzīmē, ka šie PRS modeļi ir nozīmīgs sasniegums, jo tie ir pirmie genoma mēroga modeļi, kas aptver vairāk nekā 2 000 000 *SNV*, nodrošinot ļoti plašu ģenētisku pārklājumu. Sīkāka informācija par katru izmantoto modeli ir sniegta 2.6. tabulā.

2.6. tabula

Apvienoto modeļu, kas lietoti riska aprēķinos, raksturojums

Modelis	Apraksts
modelis1	Efekta lielums, kas aprēķināts KV pacientos ar <i>BayesW</i> modeli
modelis2	Efekta lielums, kas aprēķināts KV pacientos ar <i>BayesRR-RC</i> modeli
modelis3	Efekta lielums, kas aprēķināts OV pacientos ar <i>BayesW</i> modeli
modelis4	Efekta lielums, kas aprēķināts OV pacientos ar <i>BayesRR-RC</i> modeli

KV – krūts vēzis; OV – olnīcu vēzis.

Izmantojot binomiālo loģistisko regresijas analīzi, tika novērtēta četru PRS (modelis1–4) saistība ar risku saslimt ar KV vai OV. Šī pētījuma mērķis bija noteikt nesen izstrādāto genoma mēroga PRS modeļu (*BayesW* vs. *BayesRR-RC*) efektivitāti KV un OV riska prognozēšanā reģionam specifisku *BRCA1* PV nesējās Latvijas populācijā. Tas tika panākts, salīdzinot PRS aprēķināto efektu lielumu *BRCA1* PV nesējās ar vēža diagnozi (KV un/vai OV) salīdzinājumā ar *BRCA1* PV nesējām bez vēža diagnozes (ar vēzi neskartā grupa).

Starp četriem testētajiem PRS skaidri izcēlās modelis1, kurš uzrādīja visspēcīgāko asociāciju ar KV attīstības risku. Analizētā modeļa1 OR bija 1,37 (95 % TI = 1,03–1,81, $p = 0,0291$), kā norādīts 2.7. tabulā.

**Binomiālās loģistiskās regresijas analīzes rezultāti trijās dažādās
pētījuma grupās**

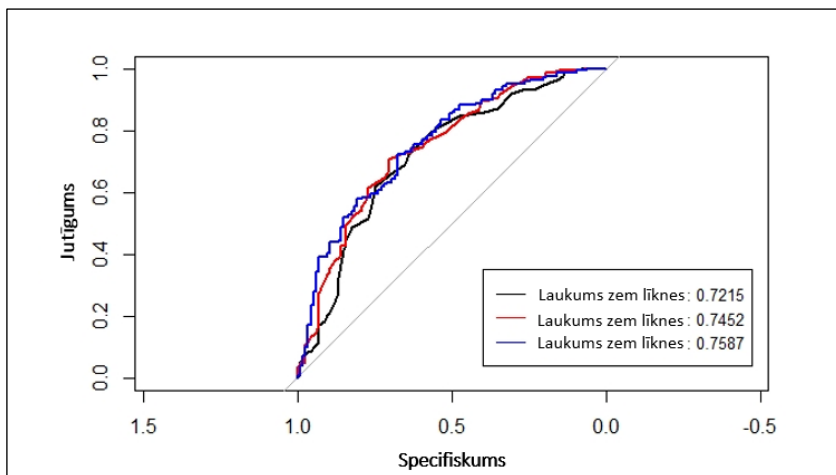
	OR	95 % TI	p vērtība
KV + OV vs. Nav vēzis			
modelis1	1,14	0,89–1,46	0,3119
modelis2	1,11	0,86–1,42	0,4205
modelis3	1,00	0,78–1,28	0,9781
modelis4	0,89	0,69–1,14	0,3514
<i>BRCA1:c.5266dup</i>	1,73	1,03–2,91	0,0375*
KV vs. Nav vēzis			
modelis1	1,37	1,03–1,81	0,0291*
modelis2	1,33	1,01–1,76	0,0423*
modelis3	1,00	0,76–1,31	0,9825
modelis4	0,95	0,72–1,25	0,7109
<i>BRCA1:c.5266dup</i>	2,55	1,44–4,53	0,0013**
OV vs. Nav vēzis			
modelis1	0,94	0,68–1,31	0,7180
modelis2	0,91	0,65–1,27	0,5800
modelis3	0,99	0,71–1,38	0,9530
modelis4	0,81	0,57–1,14	0,2250
<i>BRCA1:c.5266dup</i>	0,93	0,48–1,79	0,8170

KV – krūts vēzis; OV – olnīcu vēzis; KV + OV – apvienoti abi vēža veidi; OR – izredžu attiecība; 95 % TI – 95 % ticamības intervāls, kas noteikts PRS asociācijai ar KV un OV risku *BRCA1* PV nesējās. Riska aprēķinos tika izmantoti četri dažādi apvienotie PRS modeļi (skatīt 2.6. tabulu). * p vērtība, kas zemāka par 0,05; ** p vērtība, kas zemāka par 0,01.

Neatkarīgi no konkrētā izmantotā PRS neviens no analizētajiem modeļiem neuzrādīja statistiski nozīmīgu saistību ar OV attīstības risku ($p > 0,05$), kā parādīts 2.7. tabulā.

Tālāk tika veikta laukuma zem uztvērēja darbības raksturlīknes (*AUC*) analīze, lai novērtētu trīs dažādu riska paredzēšanas modeļu, kuros iekļauti dažādi kovariāti, tostarp PRS, prognozēšanas precizitāti (skatīt 2.3. attēlu). Īpaši jāatzīmē, ka riska paredzēšanas modelis, kas ietvēra diagnozes vecumu, vecuma kvadrātu, *BRCA1* PV un visefektīvāko PRS (modelis1), uzrādīja augstāko *AUC* vērtību 0,7587.

Veicot šo trīs riska paredzēšanas modeļu *AUC* vērtību salīdzināšanu, izmantojot ‘*bootstrap*’ metodi, mēs konstatējam statistiski nozīmīgu atšķirību, it sevišķi starp modeļi, kurā kā kovariāti bija iekļauti tikai vecums un vecuma kvadrāts, un modeļi, kurā kā kovariāti bija iekļauti vecums, vecuma kvadrāts, *BRCA1* PV un visefektīvākais PRS (modelis1) ar $p = 0,0368$.



2.3. attēls. *AUC* (laukumu zem uztvērēja operatora līknes) salīdzinājums, lai noteiktu optimālo binomiālās loģistikās regresijas analīzes modeli

Ar melnu apzīmēts modelis, kurā ir iekļauts tikai vecums un vecuma kvadrāts; ar sarkanu – modelis ar papildus pievienotu *BRCA1* PV; ar zilu – modelis ar papildus pievienotu *BRCA1* PV un visefektīvāko PRS (t.i., modelis1).

3. Diskusija

3.1. Galvenie rezultāti pētījuma kohortā

Šajā pētījumā tika analizēta saistība starp konkrētiem *BRCA1* PV (c.4035del un c.5266dup) un KV vai OV attīstību. Šo PV sadalījums pētījuma kohortā (skatīt 2.1. tabulu) ir saskaņā ar iepriekšējiem pētījumiem, apstiprinot to nozīmīgumu Latvijas populācijā (Gardovskis et al., 2005; Tikhomirova et al., 2005). Tika analizēta šo divu *BRCA1* PV penetrance un ietekme uz KV vai OV attīstības vecumu. Papildus pētījuma paraugi kalpoja par pamatu *GWAS* un *PRS* analīzei, lai identificētu ģenētiskos faktoros, kas ietekmē reģionam specifisko *BRCA1* PV penetranci. Šāda detalizēta modificējošo faktoru analīze var uzlabot izpratni par konkrētu *BRCA1* PV ietekmi uz KV un OV attīstības risku, kā arī virzīt turpmākos pētījumus, personalizētu riska novērtējumu un profilakses stratēģiju izmantojumu.

Šī pētījuma mērķis bija novērtēt specifisku *BRCA1* PV penetranci KV un OV pētījuma kohortā. Lai gan ar vēzi neskarto indivīdu vecums, kas bija mazāks nekā pārējās grupās, varētu ietekmēt penetrances aplēses, tomēr šis pētījums sniedz vērtīgu ieskatu par reģionam specifisko *BRCA1* PV penetranci.

Šī pētījuma rezultāti apstiprina konceptu, ka penetrance var atšķirties atkarībā no konkrētā *BRCA1* PV lokalizācijas. Piemēram, *BRCA1*:c.5266dup PV uzrāda lielāku penetranci KV grupā salīdzinājumā ar OV grupu, bet *BRCA1*:c.4035del PV uzrāda līdzīgu penetranci abos vēža veidos (skatīt 2.2. tabulu). Šiem PV ir novērojama genotipa–fenotipa korelācija un atšķirīga klīniskā izpausme, ko, iespējams, ietekmē PV atrašanās vieta un funkcionālā ietekme uz *BRCA1* proteīnu. Iepriekšējie pētījumi liecina, ka PV, kas atrodas *BRCA1* gēna 3' galā (piemēram, c.5266dup), ir saistīti ar augstāku risku saslimt ar KV, bet PV, kas atrodas 10. eksonā (piemēram, c.4035del), gandrīz vienlīdz

bieži ierosina KV un OV *BRCA1* PV nesējās (Milne & Antoniou, 2016; Plakhins et al., 2011).

Analizētajos datos *BRCA1:c.4035del* PV neuzrādīja statistiski nozīmīgus pierādījumus par paaugstinātu KV attīstības risku salīdzinājumā ar OV, kas apstiprina novērojumu, ka šis specifiskais *BRCA1* PV ir saistīts ar relatīvi līdzsvarotu risku abiem vēža veidiem. Tas izceļ *BRCA1* PV pozīcijas potenciālo nozīmi riska novērtēšanā (Kuchenbaecker, Hopper et al., 2017).

Tālāk tika veikta Koksas proporcionālā riska regresijas analīze pētījuma kohortā (452 indivīdi), kas atklāja statistiski nozīmīgu *BRCA1:c.5266dup* PV ietekmi uz agrāku vēža attīstību (apvienojot KV un OV grupas) salīdzinājumā ar *BRCA1:c.4035del* PV. *BRCA1:c.5266dup* PV nesējās vidējais vēža diagnozes vecums bija 46,52 gadi, savukārt *BRCA1:c.4035del* PV nesējās vidējais vēža diagnozes vecums bija 50,62 gadi (skatīt 2.1. tabulu). Riska koeficients 1,437 norāda uz 43,70 % paaugstinātu agrāka vēža attīstības risku *BRCA1:c.5266dup* PV nesējās.

KV grupā *BRCA1:c.5266dup* PV demonstrēja 56,40 % augstāku KV attīstības risku, kas atbilst iepriekš ziņotajiem datiem par Latvijas KV pacientēm (Plakhins et al., 2011). Šie rezultāti uzsver nepieciešamību pēc personalizētas pieejas ģenētiskajā konsultēšanā, iekļaujot informāciju par šiem diviem reģionam specifiskiem *BRCA1* PV riska stratēģijas izveidā, piemēram, izvēloties pastiprinātu skrīningu vai *BRRM*.

BRCA1:c.5266dup PV būtiski ietekmēja agrīnāku vēža rašanos KV grupā, bet ne OV grupā, uzsverot *BRCA1* PV specifisko ietekmi uz genotipa–fenotipa korelāciju (Milne & Antoniou, 2016; Plakhins et al., 2011). Mijiedarbība ar citiem ģenētiskiem vai vides faktoriem varētu potenciāli modificēt *BRCA1:c.5266dup* PV ietekmi uz OV attīstības risku. Iespējams, vēlīnā stadijā, kurā lielākoties tiek diagnosticēts OV, ietekmēja pētījumā reģistrēto OV diagnozes vecumu, potenciāli samazinot *BRCA1:c.5266dup* PV ietekmi uz

attīstības vecumu salīdzinājumā ar KV grupu, kurā agrīna diagnostika ir izplatītāka (Thulesius et al., 2004).

3.1.1. Uz hipotēzi balstīta *BRCA1* un *CHEK2* dubultheterozigotu analīze

Tālāk tika pētīta *CHEK2* gēna variantu ietekme uz *BRCA1* PV penetranci, jo *CHEK2* kopā ar *BRCA1* abi ir iesaistīti vienā DNS reparācijas signālceļā. *CHEK2* varianti bieži tiek novēroti KV un OV pacientiem un ir plaši pētīti vairākās Eiropas valstīs (Myszka et al., 2011; Narod & Lynch, 2007), tomēr pētījumi Latvijas populācijā ir ierobežoti. Tikai divos iepriekšējos pētījumos, ko veikuši Irmejs et al. un Plonis et al., ir analizēta konkrētu *CHEK2* variantu saistība ar KV, OV un kolorektālā vēža attīstības risku (Irmejs et al., 2006; Plonis et al., 2015).

Lai gan daudzi pētījumi liecina, ka *SNV* modificējošos gēnos ietekmē PV penetranci, pētījumi *BRCA1* un *CHEK2* dubultheterozigotās ir reti. Iepriekšējos pētījumos, galvenokārt koncentrējoties uz KV pacientiem, ir identificēti tikai 1 līdz 15 šādu dubultheterozigotu gadījumi starp tūkstošiem pacientu (Cybulski et al., 2009; Meijers-Heijboer et al., 2002; Sokolenko et al., 2014; Turnbull et al., 2012).

Atšķirībā no iepriekš veiktajiem pētījumiem, kuros galvenokārt tika salīdzināts *BRCA1* un *CHEK2* dubultheterozigotu biežums starp KV pacientiem un veselām kontrolēm no vispārējās populācijas, šī pētījuma galvenais mērķis bija izvērtēt hipotēzi, ka *CHEK2* varianti varētu ietekmēt *BRCA1* PV penetranci. Šī pētījuma mērķis bija novērtēt *BRCA1* un *CHEK2* dubultheterozigotu klātbūtni KV un OV pacientēm salīdzinājumā ar neskarto grupu, kurai uz rekrutēšanas brīdi nebija diagnosticēts vēzis, bet tās bija *BRCA1* PV nesējas.

Tika identificēti 13 *BRCA1* un *CHEK2* dubultheterozigotu gadījumi (skatīt 2.3. tabulu), kuru biežums atbilst iepriekšējiem pētījumiem. Rezultāti liecina par tendenci, ka dubultheterozigotām OV tiek diagnosticēts agrāk nekā *BRCA1* PV nesējās bez *CHEK2* varianta, tomēr šis novērojums nebija statistiski

nozīmīgs. Līdz ar to rezultāti saskan ar iepriekšējiem pētījumiem, kas neuzrāda būtisku *CHEK2* variantu ietekmi uz vēža attīstības risku *BRCA1* PV nesējās (Cybulski et al., 2009; Sokolenko et al., 2014, Sukumar et al., 2021).

Iepriekšējos pētījumos citās populācijās, kā arī Latvijā tiek novērots zemāks *CHEK2* variantu biežums *BRCA1* PV nesējās salīdzinājumā ar KV pacientiem bez *BRCA1* PV, kas liecina par negatīvu mijiedarbību starp šiem variantiem (Cybulski et al., 2009; Irmejs et al., 2006; Meijers-Heijboer et al., 2002; Plonis et al., 2015; Sokolenko et al., 2014; Turnbull et al., 2012). Šis novērojums varētu būt saistīts ar samazinātu šūnu dzīvotspēju, kurās abi gēni ir defektīvi, jo to produkti ir iesaistīti vienā DNS reparācijas signālceļā (Bartek & Lukas, 2003; Collins & Garrett, 2005; Lee et al., 2000).

Kopumā nav iegūti statistiski nozīmīgi pierādījumi par patogēno/iespējami patogēno *CHEK2* variantu ietekmi uz KV vai OV attīstības risku *BRCA1* PV nesējās. Nelielais paraugu skaits var ierobežot analīzes statistisko spēku, un ir nepieciešami pētījumi lielākās kohortās, lai palielinātu šo rezultātu ticamību.

3.1.2. Uz datiem balstīta ar vēža risku saistītu variantu identificēšana viena nukleotīda varianta (SNV) līmenī *BRCA1* PV nesējās

Lai veiktu uz datiem balstītu variantu identificēšanu *SNV* līmenī, kas saistīti ar KV vai OV attīstības risku *BRCA1* PV nesējās, tika īstenota *GWAS* analīze. Šī pētījuma mērķis bija novērtēt biežos ģenētiskos variantus, kas saistīti ar KV un OV, kā potenciālos vēža attīstības riska modifikatorus *BRCA1* PV nesējās. Tā kā pētījuma kohorta bija salīdzinoši neliela, *GWAS* statistiskais spēks bija pietiekams tikai biežo ģenētisko variantu identificēšanai.

Šajā darbā tika pētīta reģionam specifisko *BRCA1* PV (c.4035del un c.5266dup) ģenētiskie modifikatori un to saistība ar KV vai OV attīstības risku Latvijas iedzīvotāju klīniskajā kohortā. Izmantojot *GWAS* pieeju, tika identificēti

18 genomiskie riska lokusi, kas saistīti ar KV attīstības risku, un 21 genomiskais riska lokuss, kas saistīts ar OV attīstības risku. Neraugoties uz daudzām plaša mēroga *GWAS*, kas veiktas gan vispārējā populācijā, gan *BRCA1* PV nesēju vidū, kurās veiksmīgi identificēti vairāk nekā simts ar KV un OV attīstības risku saistītu lokusu, neviens no šajā pētījumā identificētajiem riska lokusiem iepriekš nav aprakstīts. Turklāt šajā kohortā neizdevās reproducēt iepriekšējo *GWAS* rezultātus (Couch et al., 2013; Kuchenbaecker et al., 2015; Milne & Antoniou, 2016; Milne et al., 2017; Yang et al., 2022).

Tas, ka šajā pētījumā netika identificēti iepriekš aprakstītie genomiskie riska lokusi, visticamāk, skaidrojams ar unikālo pētījuma dizainu un iespējamām metodoloģijas atšķirībām. Pirmkārt, lielākā daļa iepriekš identificēto ar vēža risku saistīto *SNV* tika atklāti vispārējā populācijā (Amos et al., 2017; Jurj et al., 2020; Michailidou et al., 2017; Phelan et al., 2017). Tomēr ir pierādīts, ka bieži *SNV*, kas identificēti vispārējā populācijā, var nepaaugstināt KV vai OV attīstības risku *BRCA1* PV nesējās (Coignard et al., 2021). Turklāt lielākajā daļā asociāciju pētījumu ar *BRCA1* PV nesējām ir izmantots gadījuma kontroles dizains, kur kontroles grupa sastāv no veselām sievietēm no vispārējās populācijas bez diagnosticētiem *BRCA1* PV (Milne & Antoniou, 2016). Turpretī šī pētījuma dizains specifiski koncentrējās uz *BRCA1* PV nesējām, lai varētu identificēt nesējām raksturīgus *SNV*, kas saistīti ar KV un OV attīstības risku (Coignard et al., 2021). Līdz ar to šī pētījuma rezultāti var nebūt tieši salīdzināmi ar vairākumu iepriekšējo pētījumu rezultātiem. Turklāt citi pētījumi lielākoties ir koncentrējušies uz lielu konsorciju paraugu analīzi, kas sastāv no plaša spektra *BRCA1* PV (Rebbeck et al., 2018), tomēr šajā pētījumā tika analizēti paraugi ar Latvijas populācijai un Baltijas reģionam raksturīgiem reģionam specifiskiem *BRCA1* PV (Gardovskis et al., 2005; Janavičius et al., 2014; Tamboom et al., 2010; Tikhomirova et al., 2005).

Šajā pētījumā tika identificēti 27 neatkarīgi nozīmīgi *SNV* KV grupā un 25 *SNV* OV grupā, kas pārsniedza iespējama nozīmīguma sliekšni ($p < 1 \times 10^{-6}$) un liecina par šo *SNV* potenciālo lomu vēža attīstības riska modulēšanā. Turklāt analizētie dati saturēja ievērojamu skaitu kandidāt-*SNV*, kas atrodas *LD* ($r^2 > 0,6$) ar identificētajiem neatkarīgajiem nozīmīgajiem *SNV*, kā rezultātā KV grupā tika iegūti 1152 kandidāti un OV grupā – 633 kandidāti. Lielākā daļa šo kandidātu tika identificēti genoma nekodējošos reģionos, norādot uz regulatīvo reģionu nozīmi un ietekmi uz vēža attīstības risku un uzsverot nepieciešamību pēc turpmākas padziļinātas funkcionālās izpētes. Turklāt lielākā daļa pieejamo *GWAS* rezultātu liecina, ka vairākums biežo variantu, kas saistīti ar vēža attīstības risku modificēšanu, atrodas genoma nekodējošajos reģionos, un tiek uzskatīts, ka tie ietekmē vēža risku, regulējot noteiktu gēnu ekspresiju (Amos et al., 2017; Edwards et al., 2013; Yang et al., 2022).

Nākamajās rindkopās detalizēti aplūkosim šajā pētījumā atklātos nozīmīgākos *GWAS* rezultātus, kas piedāvā vērtīgu pamatu turpmākiem pētījumiem un jaunus ieskatus sarežģītajā ģenētisko modifikatoru un reģionam specifisko *BRCA1* PV mijiedarbībā Latvijas populācijā. 2.4. tabulā ir uzskaitīti trīs pētījumā atklātie nozīmīgākie ģenētiskie varianti, kas saistīti ar KV vai OV attīstības risku. Interesanti, ka visi trīs varianti uzrādīja negatīvu efekta lielumu (beta), kas liecina par potenciālu aizsargājošu ietekmi uz vēža attīstību.

KV pacientēm statistiski nozīmīgāk saistītais vadošais variants bija introniskais variants rs2609813 *FAM107B* gēnā. Proteīnu kodējošais gēns *FAM107B*, kas pieder saimei ar sekvenču līdzību 107 (*FAM107*), joprojām ir nepietiekami izpētīts, un pieejamie bioloģiskie dati ir ierobežoti. Neraugoties uz to, šo gēnu saimes locekļu N-terminālā domēna (DUF1151) struktūra ir konservēta starp dažādām sugām, kas liecina par to lomu gēnu transkripcijas regulēšanā. *FAM107B* proteīns, iespējams, ietekmē citoskeleta pārkārtošanos, un tam ir nozīme šūnu migrācijā un proliferācijā. Tomēr molekulārie mehānismi,

kas ir *FAM107B* bioloģisko funkciju pamatā, joprojām nav skaidri izpētīti (Nakajima & Koizumi, 2014).

Balstoties uz iepriekšējiem pētījumiem, *FAM107* gēnu saime tiek uzskatīta par audzēju supresorgēnu kandidātiem, un ir novērota korelācija starp *FAM107A* gēna, kas iepriekš pazīstams kā *DRRI*, samazinātu aktivitāti un palielinātu kancerogēnēzi dažādos ļaundabīgos audzējos, tostarp nesīkšūnu plaušu vēža, nieru šūnu vēža, prostatas vēža un astrocitomas gadījumos (Liu et al., 2009; van den Boom et al., 2006; Wang et al., 2000). Lai gan *FAM107B* gēna struktūra ir līdzīga, tā specifiskā loma kā audzēju supresorgēnam joprojām nav skaidri pierādīta. Pētījumos ir novērota samazināta *FAM107B* ekspresija dažādos audzēju audos, tostarp krūts, vairogdziedzera, kuņģa un resnās zarnas vēža šūnās, kas liecina par tā potenciālu lomu audzēju attīstībā. Eksperimentāli pierādījumi liecina, ka paaugstināta *FAM107B* ekspresija inhibē vēža šūnu proliferāciju (Nakajima et al., 2010; Nakajima et al., 2012), savukārt *FAM107B* inhibīcija ievērojami palielina kuņģa vēža šūnu proliferāciju un migrāciju, kas atbalsta hipotēzi, ka *FAM107B* ir potenciāls audzēja supresorgēns (Guo et al., 2017). Turklāt samazināta *FAM107B* ekspresija ir novērota KV audos, īpaši agresīvos fenotipos, kas saistīti ar palielinātu slimības recidīvu risku un samazinātu dzīvildzi (Nakajima et al., 2012).

Aprēķinātais negatīvais efekta lielums variantam rs2609813, kas tiek paredzēts kā regulatīvā reģiona variants, liecina par tā iespējamu protektīvu efektu uz KV attīstību, potenciāli ietekmējot citu gēnu ekspresiju *BRCA1* PV nesējās. Kopumā *FAM107B* ir daudzsološs kandidātaudzēja supresorgēns KV, un turpmākajos pētījumos būtu vērtīgi koncentrēties uz šī regulatīvā reģiona varianta darbības mehānismu izpēti.

Otrs statistiski nozīmīgākais ar KV saistītais *SNV* bija rs4688094, kas atrodas iepriekš neaprašstītas *lncRNA* gēnā *RP11-384F7.1*, un tā bioloģiskā funkcija nav zināma. Tāpēc ir grūti prognozēt šā varianta funkcionālās sekas.

LncRNA ir atklātas kā svarīgi vēža attīstības un progresēšanas regulatori, kas ietekmē dažādus bioloģiskos procesus, piemēram, proliferāciju, apoptozi, metastāzes un terapijas rezistenci (Arun et al., 2018; Liu et al., 2021). Izmainīta *lncRNA* regulācija tiek saistīta ar dažādām ar vēzi saistītām īpašībām, darbojoties gan kā onkogēni, gan audzēju supresorgēni (Fonseca-Montaño et al., 2023). Tās var regulēt citu gēnu ekspresiju dažādos līmeņos, tostarp ietekmējot hromatīna modifikāciju, transkripciju un pēctranskripcijas RNS apstrādi, norādot uz to potenciālu mērķterapijas izstrādē (Gutschner & Diederichs, 2012). Turklāt izmainīta specifisku *lncRNA* regulācija ir saistīta ar dažādiem KV apakštipiem un klīniskajiem iznākumiem, atbalstot to potenciālu pielietošanu kā diagnostiskos un prognostiskos biomarķierus (Su et al., 2014; Zhao et al., 2021).

Šajā pētījumā aprakstītā *lncRNA RP11-384F7.1* iepriekš nav pētīta saistībā ar KV. Turklāt novērotais negatīvais efekta lielums liecina par iespējamu protektīvu ietekmi uz KV attīstības risku, kas saistīts ar riska alēli C. Šie rezultāti uzsver nepieciešamību turpināt pētīt rs4688094 varianta funkcionālo ietekmi un tā ietekmi uz *RP11-384F7.1* ekspresiju.

Vienīgais vadošais variants, kas OV grupā pārsniedza iespējamu nozīmīguma sliekšni, bija rs79732499, kas atrodas intergēniskā regulatīvā reģionā. Šis variants uzrāda augstu *LD* ar vairākiem *SNV C20orf194* gēnā, kas pazīstams arī kā *DNAAF9*, norādot uz iespējamu ietekmi uz *DNAAF9*.

Informācija par *DNAAF9* funkcijām ir ierobežota, taču pašreizējās zināšanas liecina, ka tas ir iesaistīts mikrocaurulīšu mijiedarbībā un tubulīna montāžas procesā (Casalou et al., 2020). Saskaņā ar *UniProt* datubāzes datiem šis gēns ir iespējams *ARL3* (ADP ribozilēšanas faktoram līdzīgā GTPāze 3) efektors. Lai gan *ARL3* funkcionālā loma vēža attīstībā joprojām nav zināma, novērojumi gliomā liecina, ka *ARL3* ir nozīme angioģenēzē un imūno šūnu infiltrācijā audzēja mikrovidē (Casalou et al., 2020; Wang et al., 2019).

Neraugoties uz identificētā rs79732499 varianta negatīvo efekta lielumu, kas liecina par iespējamu protektīvu ietekmi *BRCA1* PV nesējās, precīza *DNAAF9* funkcija OV attīstībā joprojām nav zināma. Lai izprastu tās funkciju un potenciālo ietekmi uz OV, ir nepieciešami papildu pētījumi.

Tālāk tika pētīta *SNV* ietekme uz gēnu ekspresiju krūts un olnīcu audos, izmantojot *eQTL* kartēšanu *GTEX* datos. Lai gan olnīcu audos netika novērotas nozīmīgas *SNV* gēnu asociācijas, tomēr krūts audos tika identificēti divi *SNV*, rs10178186 un rs434451, kas ietekmē gēnu ekspresiju šajos audos (skatīt 2.5. tabulu).

Pirmais identificētais *eQTL* variants rs10178186 ir saistīts ar samazinātu *ZNF514* gēna ekspresiju. Cinka pirkstu proteīni (*ZNF*), tostarp *ZNF514*, ir transkripcijas faktori, kas, kā tiek uzskatīts, ir nozīmīgi kancerogēnēzē, audzēja progresēšanā un metastāžu veidošanā dažādos vēža veidos. Neraugoties uz šo proteīnu lielo skaitu, lielākā daļa *ZNF* nav labi izpētīti (Luo et al., 2018; Ye et al., 2021).

Lai gan iepriekšējos pētījumos ir norādīts uz noteiktu *ZNF* lomu vēža attīstībā, ir nepieciešami papildu funkcionālie pētījumi, lai pilnībā izprastu to potenciālo lomu un ietekmi uz KV attīstību. Piemēram, *ZNF165* ir saistīts ar trīskārši negatīva KV (*TNBC*) attīstību, potenciāli veicinot agresīvāku kancerogēnēzi (Gibbs et al., 2020). Un pretēji, hipermetilēts *ZNF154* promoters tiek saistīts ar labākiem dzīvildzes rādītājiem rezecējama aizkuņģa dziedzera vēža gadījumā (Wiesmueller et al., 2019).

Otrais *eQTL* variants rs434451 ir saistīts ar samazinātu *SLCIA5* ekspresiju. Šis gēns kodē šūnu virsmas izšķīdušo vielu transportieri, kurš ir svarīgs neitrālu aminoskābju, it īpaši glutamīna, uzņemšanai, kas ir būtisks vēža šūnu metabolismam (Alfarsi et al., 2021; van Geldermalsen et al., 2016). Novērotā negatīvā saistība starp rs434451 variantu un *SLCIA5* ekspresiju *BRCA1* PV nesējās liecina par potenciālu lomu glutamīna metabolisma

modulēšanā un līdz ar to arī vēža attīstībā KV gadījumā. Farmakoloģiskie pētījumi ir pierādījuši, ka *SLCIA5* mediētā transporta inhibēšana samazina glutamīna uzņemšanu, kā rezultātā samazinās vēža šūnu proliferācija un palielinās šūnu nāve, īpaši *TNBC* šūnās (van Geldermalsen et al., 2016). Turklāt *SLCIA5* ekspresija ir saistīta ar jutību pret endokrīno terapiju luminālā KV gadījumā, kas liecina par tās potenciālo lietderību kā ārstēšanas atbildes reakcijas prognozējošu biomarķieri (Alfarsi et al., 2021).

Iepriekšējos pētījumos ir uzsvērtā *SLCIA5* gēna un ar to saistīto transportieru nozīme vēža šūnu metabolismā, augšanā un proliferācijā. Lai gan preklīniskajos pētījumos ir iegūti daudzsološi rezultāti ar *SLCIA5* inhibitoriem, pašlaik netiek veikti klīniskie pētījumi, kuros pārbaudītu to efektivitāti un drošību (Nachef et al., 2021). Lai pilnībā izprastu *SLCIA5* gēna lomu un tā mērķterapijas potenciālu KV gadījumā, ir nepieciešami turpmāki pētījumi.

Visbeidzot, ir nepieciešami papildu pētījumi, lai izprastu *ZNF514* un *SLCIA5* ekspresijas potenciālo nozīmi KV attīstībā *BRCA1* PV nesējās. Nepieciešami padziļinātāki pētījumi par precīziem mehānismiem, kas nosaka to nozīmi audzēja augšanā, progresēšanā un atbildes reakcijā uz terapiju.

3.1.3. Uz datiem balstīta ar vēža risku saistītu variantu identificēšana poligēnu riska variantu (PRS) līmenī *BRCA1* PV nesējās

Šajā pētījumā tika analizēta divu nesen ziņotu genoma mēroga PRS (Orliac et al., 2022), kas satur 2 174 072 *SNV*, asociācija ar KV un OV attīstības risku *BRCA1* PV nesējās. Lai gan joprojām nav skaidrs, kāda ir labākā pieeja *SNV* kopas atlasei un to efektu lielumu aprēķināšanai, lai izveidotu visefektīvāko PRS, mūsu pētījums koncentrējās uz genoma mēroga *SNV* ietekmes novērtēšanu apvienotajos PRS modeļos. Pētījuma mērķis bija noteikt, vai genoma mēroga PRS ir efektīvi riska paredzēšanā, kas varētu palielināt prognozēšanas precizitāti, salīdzinot ar parasti izmantotajām PRS izstrādes pieejām (Dareng et al., 2022).

Ņemot vērā, ka lielākā daļa PRS, tostarp šajā pētījumā analizētie, ir aprēķināti no vispārējo populāciju pētījumiem, ir svarīgi rūpīgi izvērtēt un validēt to veikspēju, jo īpaši indivīdos, kuri ir *BRCA1* PV nesēji (Jones et al., 2017; Mavaddat et al., 2019; Michailidou et al., 2017). Variablā *BRCA1* gēna PV penetrance rada būtiskas grūtības novērtēt katra indivīda vēža rašanās varbūtību, vecumu un lokalizāciju. Tādēļ ir svarīgi izpētīt efektīvas stratēģijas, lai uzsāktu profilaktisko skrīningu un klīnisku riska samazināšanu augsta riska sievietēm (Chen et al., 2020; Downs et al., 2019). PRS ir potenciāls stratificēt indivīdus, pamatojoties uz viņu slimības risku (Mars et al., 2020). Tomēr, lai sasniegtu šo mērķi un integrētu PRS klīniskajā praksē, ir būtiski noteikt optimālo *SNV* kopumu, kas veido visefektīvāko PRS.

Šī pētījuma rezultāti demonstrē vislabākā *BayesW* PRS efektivitāti, lai prognozētu indivīda KV attīstības risku. Lai gan *BayesRR-RC* PRS arī labi prognozēja KV attīstības risku, tomēr *BayesW* PRS bija efektīvāks (skatīt 2.7. tabulu).

Iepriekšējā pētījumā *Kuchenbaecker* et al. izstrādāja trīs PRS KV (vispārējam KV, ER pozitīvam un ER negatīvam KV), kā arī vienu PRS OV. Viņu pētījumā tika izmantoti dati par 15 252 *BRCA1* PV nesējām, atklājot statistiski nozīmīgu saistību starp PRS un risku saslimt gan ar KV, gan OV. Visnozīmīgākā saistība ar KV attīstības risku bija PRS, kas izveidots ER negatīvam KV ($HR = 1,27$, 95 % $TI = 1,23-1,31$, $p = 8,2 \times 10^{-53}$) (*Kuchenbaecker, McGuffog et al., 2017*).

Līdzīgus novērojumus atkārtoja arī *Barnes* et al. pētījumā, kurā bija iekļautas 9473 *BRCA1* PV nesējas ar diagnosticētu KV (*Barnes et al., 2020*), demonstrējot, ka ER negatīvs PRS uzrādīja visspēcīgāko saistību ar KV attīstības risku *BRCA1* PV nesējās ($HR = 1,29$, 95 % $TI = 1,25-1,33$, $p = 3 \times 10^{-72}$). Ņemot vērā, ka ER negatīvs KV ir dominējošais audzēja apakštips *BRCA1* PV nesējās (*Foulkes et al., 2004*), šie pētījumi uzsver KV apakštipam specifisku PRS saistību

ar KV attīstības risku. Šie rezultāti pierāda, ka visefektīvākā KV attīstības riska prognozēšana ietver visaptverošas klīniskas informācijas integrēšanu analizē (Barnes et al., 2020; Kuchenbaecker, McGuffog et al., 2017). Diemžēl nepietiekamu klīnisko datu dēļ šajā promocijas darbā nebija iespējams iekļaut informāciju par ER statusu, kas bija pieejama tikai nelielai daļai (< 80) KV pacienšu.

Citā pētījumā *Mavaddat* et al. pierādīja spēcīgu saistību starp PRS un KV attīstības risku vispārējā populācijā (OR = 1,61, 95 % TI = 1,57–1,65, ar $AUC = 0,630$, 95 % TI = 0,628–0,651) (*Mavaddat* et al., 2019). Mūsu rezultāti atbilst iepriekš veiktajiem pētījumiem, norādot, ka aprēķinātā KV OR indivīdiem ar *BRCA1* PV ir zemāka nekā iepriekš publicētiem aprēķiniem vispārējā populācijā. Tas liecina, ka PRS ietvaros pastāv potenciāla *SNV* apakšgrupa, kuru efekts, iespējams, nekombinējas multiplikatīvi ar *BRCA1* PV statusu. Tomēr ir būtiski pieminēt, ka tiešo salīdzinājumu potenciālie ierobežojumi var rasties no atšķirībām pētījumu dizainā un kohortas izmērā (*Kuchenbaecker, McGuffog* et al., 2017).

Atšķirībā no iepriekšējiem pētījumiem, kas konsekventi norādīja uz būtisku saistību starp PRS un OV attīstības risku, mūsu pētījumā šāda saistība netika novērota (*Barnes* et al., 2020; *Kuchenbaecker, McGuffog* et al., 2017). Pētījuma dati liecina, ka genoma mēroga PRS bija efektīvāks, lai prognozētu risku saslimt ar KV nekā ar OV *BRCA1* PV nesējās (OR = 1,37, 95 % TI = 1,03–1,81, $p = 0,029$ KV grupā vs. OR = 0,99, 95 % TI = 0,71–1,38, $p = 0,95$ OV grupā). Novērotos rezultātus, iespējams, ietekmēja ierobežotais kohortas izmērs – 121 *BRCA1* PV nesējas ar diagnosticētu OV.

3.2. Pētījuma priekšrocības un ierobežojumi

Uz 2023. gada novembri *NCBI ClinVar* datubāze (<https://www.ncbi.nlm.nih.gov/clinvar/>) saturēja 3264 pārmantojamu *BRCA1* PV ierakstus (tostarp delēcijas, duplikācijas, indeli, insercijas un *SNV*, visi < 50 bp). Tikmēr šis promocijas darbs koncentrējās uz ģenētiski homogēnu kohortu, kas sastāvēja no sievietēm, kurām bija identificēts viens no diviem reģionam specifiskiem *BRCA1* PV (c.4035del vai c.5266dup). Lielākā daļa pašreizējo penetrances aplēšu ir iegūtas liela mēroga pētījumos, kuros ir apvienoti dati no vairākām populācijām, bieži vien neņemot vērā penetranci, kas raksturīga PV ar ciltstēva efektu atsevišķās populācijās. Populācijai raksturīgā ģenētiskā struktūra var ietekmēt pētījumu rezultātus, jo noteikti modificējošo gēnu *SNV* var būt biežāk sastopami vienā populācijā, bet reti citā (Narod, 2002). To apstiprina *Pankratov et al.* pētījums, kurā ierosināja, ka vietējās populācijas vēsture un ģenētiskā struktūra var būtiski ietekmēt asociāciju analīzes rezultātus (Pankratov et al., 2020). Tāpēc pētījumi populācijās, kurās ir novērojams ciltstēva efekts, ir lietderīgi, lai novērtētu reģionam specifisko *BRCA1* PV penetranci un noteiktu modificējošos ģenētiskos faktoros.

Papildus tam mūsu rezultāti ir neatkarīgs nesen izveidoto genoma mēroga PRS novērtējums tieši Latvijas populācijas reģionam specifisko *BRCA1* PV nesējās, lai gan izmantotie PRS sākotnēji tika izstrādāti vispārējā populācijā, izmantojot Apvienotās Karalistes Biobankas un Igaunijas Genoma centra dalībnieku datus (Orliac et al., 2022). Mēs uzskatām, ka šiem genoma mēroga PRS ir potenciāls nodrošināt līdzvērtīgas vai pat labākas vēža riska prognozēšanas iespējas salīdzinājumā ar iepriekš izstrādātajiem PRS.

Tomēr šajā pētījumā mēs saskārāmies arī ar vairākiem ierobežojumiem, pirmkārt, sakarā ar nelielo indivīdu skaitu ar dubultheterozigotisku *BRCA1* un *CHEK2* genotipu. Tas galvenokārt bija saistīts ar ierobežoto kohortas lielumu un šī genotipa retumu (Cybulski et al., 2009). Lai pārbaudītu hipotēzi, ka

dubultheterozigotām var būt lielāks KV vai OV attīstības risks, ir nepieciešama lielāka kohorta, ko potenciāli var panākt ar konsorciju datiem vai lielākiem sadarbības pētījumiem. Papildu dubultheterozigotas varētu tikt identificētas, pieaugot visa eksoma sekvencēšanas (*WES*) vai *WGS* pielietojumam.

Citi ierobežojumi ietver nelielo KV un OV gadījumu skaitu ar identificētu pārmantojamu *BRCA1* PV un iespējamās atlases novirzes, jo paraugi ir klīniskā kohorta, kas atlasīta pēc diagnostiskās pārmantoto PV testēšanas. Pētījuma koncentrēšanās uz diviem konkrētiem *BRCA1* PV var pilnībā neatspoguļot visu *BRCA1* PV nesēju populāciju. Papildus mūsu PRS analīzi ierobežoja trūkstošie *SNV*, ko varētu uzlabot, imputācijas posmā izmantojot populācijai specifisku references paneli no Latvijas populācijas genoma datubāzes (LGDB), tiklīdz būs pieejami attiecīgie *WGS* dati (Rovite et al., 2018).

Būtiskā vecuma atšķirība starp analizētajām grupām, iespējams, ietekmēja penetrances aplēses, ko varētu novērst turpmākajos pētījumos, izmantojot vecuma saskaņotu pētījuma kohortu. Turklāt detalizētas klīniskās informācijas trūkums par audzēju fenotipiem padarīja šī pētījuma rezultātus par vidējo novērtējumu starp visiem KV un OV fenotipiem.

Visbeidzot, sarežģītā *SNV* mijiedarbība un identificēto nekodējošo variantu pārsvars izceļ kritisku nepieciešamību pēc plašām pēc-*GWAS* analīzēm un padziļinātiem funkcionāliem pētījumiem, lai interpretētu identificēto *SNV* ietekmi uz konkrētiem fenotipiem (Milne & Antoniou, 2016; Yang et al., 2022).

3.3. Nākotnes perspektīvas

Izpratne par to, kā identificētie ģenētiskie varianti ietekmē konkrētu *BRCA1* PV (c.4035del un c.5266dup) penetranci, ir ļoti svarīga, lai precīzāk novērtētu risku un izstrādātu potenciālās profilaktiskās un terapeitiskās stratēģijas indivīdiem, kuri ir šo *BRCA1* PV nesēji (Mars et al., 2020).

Viens no nākotnes virzieniem varētu ietvert longitudinālus pētījumus, kuros dalībnieki tiek rekrutēti jaunākā vecumā un novēroti ilgāku laika periodu. Turklāt, iekļaujot plašākus datus par modificējamiem riska faktoriem, tostarp smēķēšanu, alkohola patēriņu, fiziskām aktivitātēm un uztura paradumiem, tiktu uzlabota riska novērtēšana (Milne & Antoniou, 2016). Cits perspektīvs virziens varētu būt dažādu tehnoloģiju, piemēram, *WGS*, izmantošana, kas potenciāli var atklāt papildu variantus, kurus nav iespējams noteikt, izmantojot hromosomu mikrorindu analīzes sistēmu. Tā, piemēram, *LGDB* ir daudzsološa iniciatīva, ar kuras palīdzību varētu pētīt *BRCA1* PV Latvijas populācijā (Rovite et al., 2018), kā arī citu to penetranci ietekmējošu ģenētisko faktoru, tostarp *PRS*, izpēti. Šī iniciatīva, izmantojot *WGS* datus, palielina iespēju atklāt vairāk klīniski nozīmīgus variantus.

Nākotnē ir nepieciešama šajā pētījumā atklāto rezultātu validācija, izmantojot lielāku pētījuma kohortu, kas sastāv no reģionam specifisko *BRCA1* PV nesējām, un šis pētījums var kalpot kā preliminārie dati plašākam visu pieejamo *PRS* salīdzinājumam. Ir svarīgi uzsvērt, ka šajā analīzē netika ņemti vērā sekundāro ļaundabīgo audzēju riski, koncentrējoties tikai uz primāru *KV* vai *OV* diagnozi. Nākotnē būtu lietderīgi izpētīt, vai testētie *PRS* palīdz prognozēt sekundāro vēžu risku *BRCA1* PV nesējās.

Secinājumi

1. Balstoties uz veikto pētījumu, neviens no analizētajiem *CHEK2* variantiem būtiski neietekmē *BRCA1* gēna patogēno variantu (c.4035del un c.5266dup) penetranci.
2. Veiktajā pētījumā krūts vēža pacientu grupā introniskais variants rs2609813 *FAM107B* gēnā ir statistiski visbūtiskāk saistīts ar *BRCA1* patogēno variantu penetranci ($p = 2,33 \times 10^{-7}$, OR = 0,28).
3. Veiktajā pētījumā olnīcu vēža pacientu grupā variants rs79732499, kas atrodas genoma nekodējošā un, iespējams, regulējošā reģionā, ir statistiski visbūtiskāk saistīts ar *BRCA1* patogēno variantu penetranci ($p = 1,38 \times 10^{-7}$, OR = 0,00031).
4. Veiktajā pētījumā analizēto genoma mēroga PRS modeļu vidū *BayesW* PRS modelis vislabāk palīdz novērtēt krūts vēža attīstības risku *BRCA1* patogēno variantu (c.4035del vai c.5266dup) nesējās, un tā izmantošana nākotnē varētu uzlabot krūts vēža pacientu riska grupu stratifikāciju un profilakses stratēģiju izvēli *BRCA1* patogēno variantu nesējās.

Priekšlikumi

Ņemot vērā LGDB iniciatīvas attīstību un Latvijas donoru skaita pieaugumu, mēs ierosinām sistemātisku reģionam specifisko *BRCA1* PV skrīningu Latvijas populācijā, izmantojot pieeju, kas balstīta uz primāru genotipa noteikšanu. Šis varētu būt progresīvs solis ceļā uz personalizētāku un efektīvāku veselības aprūpes sistēmu, kuru iedvesmojušas vairāku citu globālo biobanku projektu veiksmīgās pieredzes, piemēram, Igaunijā vai Austrālijā (Leitsalu et al., 2021; Rowley et al., 2019). Pieveja, kas orientēta uz primāru genotipa noteikšanu, piedāvā inovatīvu veidu, kā identificēt indivīdus ar klīniski nozīmīgiem PV augstas penetrances *BRCA1* gēnā neatkarīgi no viņu ģimenes anamnēzes vai medicīniskām indikācijām. Papildus šāda pieeja ļautu veikt vēža riska stratifikāciju, balstoties uz personalizētiem PRS.

Minētā priekšlikuma galvenais mērķis ir uzlabot riska stratifikāciju un ilgtermiņa rezultātus reģionam specifisko *BRCA1* PV nesējos Latvijas populācijā, kuri, iespējams, nav informēti par savu ģenētisko predispozīciju uz KV vai OV attīstību. Uzlabota riska stratifikācijas stratēģija dotu iespēju indivīdiem un veselības aprūpes sniedzējiem pieņemt lēmumu par mērķtiecīgākiem un efektīvākiem profilakses pasākumiem, potenciāli samazinot KV un OV saslimstību un ietekmi uz sabiedrības veselību.

Indivīdiem, kuriem identificēti klīniski nozīmīgi PV vai augsta riska PRS rādītāji, tiktu nodrošināta plaša ģenētiskā konsultēšana par vēža risku. Konkrētā stratēģija varētu uzlabot augsta riska indivīdu un viņu radnieku ilgtermiņa rezultātus, prioritizējot ģenētisko skrīningu un atkārtotu sazināšanos ar indivīdiem, kuriem tiktu atklāti pārmantojami klīniski nozīmīgi PV, tādējādi veicinot kopējo Latvijas iedzīvotāju veselību.

Publikāciju, ziņojumu un patentu saraksts par promocijas darba tēmu

Publikācijas:

1. **Berga-Švītiņa, E.**, Pirsko, V., Nakazawa-Miklaševiča, M., Maksimenko, J., Gardovskis, J., and Miklaševičs, E. 2023. Penetrance of *CHEK2* and *BRCA1* Double Heterozygotes in Breast and/or Ovarian Cancer Patients. Proceedings of the Latvian Academy of Sciences. Section B. Natural, Exact, and Applied Sciences; vol.77, no.2, 137–140. doi: 10.2478/prolas-2023-0020.
2. **Berga-Švītiņa, E.**, Maksimenko, J., Miklaševičs, E., Fischer, K., Vilne, B., Māgi, R. 2023. Polygenic Risk Score Predicts Modified Risk in *BRCA1* Pathogenic Variant c.4035del and c.5266dup Carriers in Breast Cancer Patients. *Cancers (Basel)*; 15(11):2957. doi: 10.3390/cancers15112957.

Ziņojumi un tēzes:

1. **Berga-Švītiņa, E.**, Pirsko, V., Nakazawa-Miklaševiča, M., Maksimenko, J., Gardovskis, J., and Miklaševičs, E. 2019. *CHEK2* Pathogenic variants do not Change Penetrance of *BRCA1* variants c.4034delA and c.5266dupC. Poster presentation at Rīga Stradiņš University International Research Conference on Medical and Health Care Sciences “Knowledge for Use in Practice”: Abstracts, 1.–3.04.2019, 70.
2. **Berga-Švītiņa, E.**, Pirsko, V., Nakazawa-Miklaševiča, M., Maksimenko, J., Gardovskis, J., and Miklaševičs, E. 2023. Identifying Genetic Factors Associated with Breast or Ovarian Cancer Risk in *BRCA1* Pathogenic Variant Carriers. Oral presentation at Rīga Stradiņš University International Research Conference on Medical and Health Care Sciences “Knowledge for Use in Practice”: Abstracts, 29.–31.03.2023.
3. **Berga-Švītiņa, E.**, Loža, P., Maksimenko, J., Fischer, K., Miklaševičs, E., Māgi, R., and Vilne, B. 2023. Exploring genetic factors and polygenic risk scores to predict breast and ovarian cancer risk in *BRCA1* pathogenic variant carriers. Oral presentations at Precision Medicine Networking Forum, 12.–13.10.2023.

Literatūras un avotu saraksts

1. Alfarsi, L. H., El Ansari, R., Craze, M. L., Mohammed, O. J., Masisi, B. K., Ellis, I. O., . . . Green, A. R. (2021). SLC1A5 co-expression with TALDO1 associates with endocrine therapy failure in estrogen receptor-positive breast cancer. *Breast Cancer Res Treat*, *189*(2), 317–331. <https://doi.org/10.1007/s10549-021-06298-1>
2. Amos, C. I., Dennis, J., Wang, Z., Byun, J., Schumacher, F. R., Gayther, S. A., . . . Easton, D. F. (2017). The OncoArray Consortium: A Network for Understanding the Genetic Architecture of Common Cancers. *Cancer Epidemiol Biomarkers Prev*, *26*(1), 126–135. <https://doi.org/10.1158/1055-9965.EPI-16-0106>
3. Angeli, D., Salvi, S., & Tedaldi, G. (2020). Genetic Predisposition to Breast and Ovarian Cancers: How Many and Which Genes to Test? *Int J Mol Sci*, *21*(3). <https://doi.org/10.3390/ijms21031128>
4. Arun, G., Diermeier, S. D., & Spector, D. L. (2018). Therapeutic Targeting of Long Non-Coding RNAs in Cancer. *Trends Mol Med*, *24*(3), 257–277. <https://doi.org/10.1016/j.molmed.2018.01.001>
5. Barnes, D. R., Rookus, M. A., McGuffog, L., Leslie, G., Mooij, T. M., Dennis, J., . . . CIMBA (2020). Polygenic risk scores and breast and epithelial ovarian cancer risks for carriers of BRCA1 and BRCA2 pathogenic variants. *Genet Med*, *22*(10), 1653–1666. <https://doi.org/10.1038/s41436-020-0862-x>
6. Bartek, J., & Lukas, J. (2003). Chk1 and Chk2 kinases in checkpoint control and cancer. *Cancer Cell*, *3*(5), 421–429. [https://doi.org/10.1016/s1535-6108\(03\)00110-7](https://doi.org/10.1016/s1535-6108(03)00110-7)
7. Borde, J., Laitman, Y., Blümcke, B., Niederacher, D., Weber-Lassalle, K., Sutter, C., . . . Ernst, C. (2022). Polygenic risk scores indicate extreme ages at onset of breast cancer in female BRCA1/2 pathogenic variant carriers. *BMC Cancer*, *22*(1), 706. <https://doi.org/10.1186/s12885-022-09780-1>
8. Browning, B. L., Zhou, Y., & Browning, S. R. (2018). A One-Penny Imputed Genome from Next-Generation Reference Panels. *Am J Hum Genet*, *103*(3), 338–348. <https://doi.org/10.1016/j.ajhg.2018.07.015>
9. Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., . . . Marchini, J. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature*, *562*(7726), 203–209. <https://doi.org/10.1038/s41586-018-0579-z>
10. Carstensen. (2022). *Epi: A Package for Statistical Analysis in Epidemiology*. Oxford University Press: Oxford, UK. Available online: <https://CRAN.R-project.org/package=Epi> (accessed on 22 May 2023).
11. Casalou, C., Ferreira, A., & Barral, D. C. (2020). The Role of ARF Family Proteins and Their Regulators and Effectors in Cancer Progression: A Therapeutic Perspective. *Front Cell Dev Biol*, *8*, 217. <https://doi.org/10.3389/fcell.2020.00217>

12. CDC. (2020). Statistical Data. Statistical Data. Centre for Disease Prevention and Control. Available online: <https://www.spkc.gov.lv/lv/statistikas-dati> (accessed on 3 March 2023).
13. Chen, H., Wang, C., Conomos, M. P., Stilp, A. M., Li, Z., Sofer, T., . . . Lin, X. (2016). Control for Population Structure and Relatedness for Binary Traits in Genetic Association Studies via Logistic Mixed Models. *Am J Hum Genet*, *98*(4), 653–666. <https://doi.org/10.1016/j.ajhg.2016.02.012>
14. Chen, J., Bae, E., Zhang, L., Hughes, K., Parmigiani, G., Braun, D., & Rebbeck, T. R. (2020). Penetrance of Breast and Ovarian Cancer in Women Who Carry a. *JNCI Cancer Spectr*, *4*(4), pkaa029. <https://doi.org/10.1093/jncics/pkaa029>
15. Coignard, J., Lush, M., Beesley, J., O'Mara, T. A., Dennis, J., Tyrer, J. P., . . . Antoniou, A. C. (2021). A case-only study to identify genetic modifiers of breast cancer risk for BRCA1/BRCA2 mutation carriers. *Nat Commun*, *12*(1), 1078. <https://doi.org/10.1038/s41467-020-20496-3>
16. Collins, I., & Garrett, M. D. (2005). Targeting the cell division cycle in cancer: CDK and cell cycle checkpoint kinase inhibitors. *Curr Opin Pharmacol*, *5*(4), 366–373. <https://doi.org/10.1016/j.coph.2005.04.009>
17. Cooper, D. N., Krawczak, M., Polychronakos, C., Tyler-Smith, C., & Kehrer-Sawatzki, H. (2013). Where genotype is not predictive of phenotype: towards an understanding of the molecular basis of reduced penetrance in human inherited disease. *Hum Genet*, *132*(10), 1077–1130. <https://doi.org/10.1007/s00439-013-1331-2>
18. Couch, F. J., Wang, X., McGuffog, L., Lee, A., Olszow, C., Kuchenbaecker, K. B., . . . CIMBA. (2013). Genome-wide association study in BRCA1 mutation carriers identifies novel loci associated with breast and ovarian cancer risk. *PLoS Genet*, *9*(3), e1003212. <https://doi.org/10.1371/journal.pgen.1003212>
19. Cybulski, C., Górski, B., Huzarski, T., Byrski, T., Gronwald, J., Debniak, T., . . . Lubinski, J. (2009). Effect of CHEK2 missense variant I157T on the risk of breast cancer in carriers of other CHEK2 or BRCA1 mutations. *J Med Genet*, *46*(2), 132–135. <https://doi.org/10.1136/jmg.2008.061697>
20. Cybulski, C., Górski, B., Huzarski, T., Masojć, B., Mierzejewski, M., Debniak, T., . . . Lubiński, J. (2004). CHEK2 is a multiorgan cancer susceptibility gene. *Am J Hum Genet*, *75*(6), 1131–1135. <https://doi.org/10.1086/426403>
21. Cybulski, C., Wokołorczyk, D., Huzarski, T., Byrski, T., Gronwald, J., Górski, B., . . . Lubiński, J. (2006). A large germline deletion in the Chek2 kinase gene is associated with an increased risk of prostate cancer. *J Med Genet*, *43*(11), 863–866. <https://doi.org/10.1136/jmg.2006.044974>
22. Cybulski, C., Wokołorczyk, D., Kładny, J., Kurzawski, G., Suchy, J., Grabowska, E., . . . Lubiński, J. (2007). Germline CHEK2 mutations and colorectal cancer risk: different effects of a missense and truncating mutations? *Eur J Hum Genet*, *15*(2), 237–241. <https://doi.org/10.1038/sj.ejhg.5201734>

23. Dareng, E. O., Tyrer, J. P., Barnes, D. R., Jones, M. R., Yang, X., Aben, K. K. H., . . . Pharoah, P. D. P. (2022). Polygenic risk modeling for prediction of epithelial ovarian cancer risk. *Eur J Hum Genet*, *30*(3), 349–362. <https://doi.org/10.1038/s41431-021-00987-7>
24. Downs, B., Sherman, S., Cui, J., Kim, Y. C., Snyder, C., Christensen, M., . . . Wang, S. M. (2019). Common genetic variants contribute to incomplete penetrance: evidence from cancer-free BRCA1 mutation carriers. *Eur J Cancer*, *107*, 68–78. <https://doi.org/10.1016/j.ejca.2018.10.022>
25. Edwards, S. L., Beesley, J., French, J. D., & Dunning, A. M. (2013). Beyond GWASs: illuminating the dark road from association to function. *Am J Hum Genet*, *93*(5), 779–797. <https://doi.org/10.1016/j.ajhg.2013.10.012>
26. Fonseca-Montaño, M. A., Vázquez-Santillán, K. I., & Hidalgo-Miranda, A. (2023). The current advances of lncRNAs in breast cancer immunobiology research. *Front Immunol*, *14*, 1194300. <https://doi.org/10.3389/fimmu.2023.1194300>
27. Foulkes, W. D., Metcalfe, K., Sun, P., Hanna, W. M., Lynch, H. T., Ghadirian, P., . . . Narod, S. A. (2004). Estrogen receptor status in BRCA1- and BRCA2-related breast cancer: the influence of age, grade, and histological type. *Clin Cancer Res*, *10*(6), 2029–2034. <https://doi.org/10.1158/1078-0432.ccr-03-1061>
28. Gardovskis, A., Irmejs, A., Miklasevics, E., Borosenko, V., Bitina, M., Melbarde-Gorkusa, I., . . . Gardovskis, J. (2005). Clinical, molecular and geographical features of hereditary breast/ovarian cancer in latvia. *Hered Cancer Clin Pract*, *3*(2), 71–76. <https://doi.org/10.1186/1897-4287-3-2-71>
29. Gibbs, Z. A., Reza, L. C., Cheng, C. C., Westcott, J. M., McGlynn, K., & Whitehurst, A. W. (2020). The testis protein ZNF165 is a SMAD3 cofactor that coordinates oncogenic TGFβ signaling in triple-negative breast cancer. *Elife*, *9*. <https://doi.org/10.7554/eLife.57679>
30. Guo, J., Bian, Y., Wang, Y., Chen, L., Yu, A., & Sun, X. (2017). FAM107B is regulated by S100A4 and mediates the effect of S100A4 on the proliferation and migration of MGC803 gastric cancer cells. *Cell Biol Int*, *41*(10), 1103–1109. <https://doi.org/10.1002/cbin.10816>
31. Guo, Q., Schmidt, M. K., Kraft, P., Canisius, S., Chen, C., Khan, S., . . . Pharoah, P. D. P. (2015). Identification of novel genetic markers of breast cancer survival. *J Natl Cancer Inst*, *107*(5). <https://doi.org/10.1093/jnci/djv081>
32. Guo, Y., He, J., Zhao, S., Wu, H., Zhong, X., Sheng, Q., . . . Long, J. (2014). Illumina human exome genotyping array clustering and quality control. *Nat Protoc*, *9*(11), 2643–2662. <https://doi.org/10.1038/nprot.2014.174>
33. Gutschner, T., & Diederichs, S. (2012). The hallmarks of cancer: a long non-coding RNA point of view. *RNA Biol*, *9*(6), 703–719. <https://doi.org/10.4161/rna.20481>

34. Irmejs, A., Miklasevics, E., Boroschenko, V., Gardovskis, A., Vanags, A., Melbarde-Gorkusa, I., . . . Gardovskis, J. (2006). Pilot study on low penetrance breast and colorectal cancer predisposition markers in latvia. *Hered Cancer Clin Pract*, 4(1), 48–51. <https://doi.org/10.1186/1897-4287-4-1-48>
35. Janavičius, R., Rudaitis, V., Mickys, U., Elsakov, P., & Griškevičius, L. (2014). Comprehensive BRCA1 and BRCA2 mutational profile in Lithuania. *Cancer Genet*, 207(5), 195–205. <https://doi.org/10.1016/j.cancergen.2014.05.002>
36. Jones, M. R., Kamara, D., Karlan, B. Y., Pharoah, P. D. P., & Gayther, S. A. (2017). Genetic epidemiology of ovarian cancer and prospects for polygenic risk prediction. *Gynecol Oncol*, 147(3), 705–713. <https://doi.org/10.1016/j.ygyno.2017.10.001>
37. Jurj, M. A., Buse, M., Zimta, A. A., Paradiso, A., Korban, S. S., Pop, L. A., & Berindan-Neagoe, I. (2020). Critical Analysis of Genome-Wide Association Studies: Triple Negative Breast Cancer. *Int J Mol Sci*, 21(16). <https://doi.org/10.3390/ijms21165835>
38. Jürgens, H., Roht, L., Leitsalu, L., Nõukas, M., Palover, M., Nikopensusius, T., . . . Tõnisson, N. (2022). Precise, Genotype-First Breast Cancer Prevention: Experience With Transferring Monogenic Findings From a Population Biobank to the Clinical Setting. *Front Genet*, 13, 881100. <https://doi.org/10.3389/fgene.2022.881100>
39. Kuchenbaecker, K. B., Hopper, J. L., Barnes, D. R., Phillips, K. A., Mooij, T. M., Roos-Blom, M. J., . . . Olsson, H. (2017). Risks of Breast, Ovarian, and Contralateral Breast Cancer for BRCA1 and BRCA2 Mutation Carriers. *JAMA*, 317(23), 2402–2416. <https://doi.org/10.1001/jama.2017.7112>
40. Kuchenbaecker, K. B., McGuffog, L., Barrowdale, D., Lee, A., Soucy, P., Dennis, J., . . . Antoniou, A. C. (2017). Evaluation of Polygenic Risk Scores for Breast and Ovarian Cancer Risk Prediction in BRCA1 and BRCA2 Mutation Carriers. *J Natl Cancer Inst*, 109(7). <https://doi.org/10.1093/jnci/djw302>
41. Kuchenbaecker, K. B., Ramus, S. J., Tyrer, J., Lee, A., Shen, H. C., Beesley, J., . . . BRCA2, CIMBA (2015). Identification of six new susceptibility loci for invasive epithelial ovarian cancer. *Nat Genet*, 47(2), 164–171. <https://doi.org/10.1038/ng.3185>
42. Lee, J. S., Collins, K. M., Brown, A. L., Lee, C. H., & Chung, J. H. (2000). hCds1-mediated phosphorylation of BRCA1 regulates the DNA damage response. *Nature*, 404(6774), 201–204. <https://doi.org/10.1038/35004614>
43. Leitsalu, L., Palover, M., Sikka, T. T., Reigo, A., Kals, M., Pärn, K., . . . Tõnisson, N. (2021). Genotype-first approach to the detection of hereditary breast and ovarian cancer risk, and effects of risk disclosure to biobank participants. *Eur J Hum Genet*, 29(3), 471–481. <https://doi.org/10.1038/s41431-020-00760-2>
44. Liu, Q., Zhao, X. Y., Bai, R. Z., Liang, S. F., Nie, C. L., Yuan, Z., . . . Wei, Y. Q. (2009). Induction of tumor inhibition and apoptosis by a candidate tumor suppressor gene DRR1 on 3p21.1. *Oncol Rep*, 22(5), 1069–1075.

45. Liu, S., Sun, Y., Hou, Y., Yang, L., Wan, X., Qin, Y., . . . Liu, M. (2021). A novel lncRNA ROPM-mediated lipid metabolism governs breast cancer stem cell properties. *J Hematol Oncol*, *14*(1), 178. <https://doi.org/10.1186/s13045-021-01194-z>
46. Loh, P. R., Palamara, P. F., & Price, A. L. (2016). Fast and accurate long-range phasing in a UK Biobank cohort. *Nat Genet*, *48*(7), 811–816. <https://doi.org/10.1038/ng.3571>
47. Luo, A., Zhang, K., Zhao, Y., Zhu, Z., Fu, L., & Dong, J. T. (2018). ZNF121 interacts with ZBRK1 and BRCA1 to regulate their target genes in mammary epithelial cells. *FEBS Open Bio*, *8*(12), 1943–1952. <https://doi.org/10.1002/2211-5463.12530>
48. Mars, N., Widén, E., Kerminen, S., Meretoja, T., Pirinen, M., Della Briotta Parolo, P., . . . Ripatti, S. (2020). The role of polygenic risk and susceptibility genes in breast cancer over the course of life. *Nat Commun*, *11*(1), 6383. <https://doi.org/10.1038/s41467-020-19966-5>
49. Mavaddat, N., Michailidou, K., Dennis, J., Lush, M., Fachal, L., Lee, A., . . . Easton, D. F. (2019). Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. *Am J Hum Genet*, *104*(1), 21–34. <https://doi.org/10.1016/j.ajhg.2018.11.002>
50. Meijers-Heijboer, H., van den Ouweland, A., Klijn, J., Wasielewski, M., de Snoo, A., Oldenburg, R., . . . CHEK2-Breast Cancer Consortium (2002). Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nat Genet*, *31*(1), 55–59. <https://doi.org/10.1038/ng879>
51. Michailidou, K., Beesley, J., Lindstrom, S., Canisius, S., Dennis, J., Lush, M. J., . . . Easton, D. F. (2015). Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. *Nat Genet*, *47*(4), 373–380. <https://doi.org/10.1038/ng.3242>
52. Michailidou, K., Hall, P., Gonzalez-Neira, A., Ghoussaini, M., Dennis, J., Milne, R. L., . . . Easton, D. F. (2013). Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat Genet*, *45*(4), 353–361, 361e351–352. <https://doi.org/10.1038/ng.2563>
53. Michailidou, K., Lindström, S., Dennis, J., Beesley, J., Hui, S., Kar, S., . . . Easton, D. F. (2017). Association analysis identifies 65 new breast cancer risk loci. *Nature*, *551*(7678), 92–94. <https://doi.org/10.1038/nature24284>
54. Milne, R. L., & Antoniou, A. C. (2016). Modifiers of breast and ovarian cancer risks for BRCA1 and BRCA2 mutation carriers. *Endocr Relat Cancer*, *23*(10), T69–84. <https://doi.org/10.1530/ERC-16-0277>
55. Milne, R. L., Kuchenbaecker, K. B., Michailidou, K., Beesley, J., Kar, S., Lindström, S., . . . Simard, J. (2017). Identification of ten variants associated with risk of estrogen-receptor-negative breast cancer. *Nat Genet*, *49*(12), 1767–1778. <https://doi.org/10.1038/ng.3785>

56. Mitt, M., Kals, M., Pärn, K., Gabriel, S. B., Lander, E. S., Palotie, A., . . . 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. <https://doi.org/10.1038/ejhg.2017.51>
57. Myszka, A., Karpinski, P., Slezak, R., Czernarmazowicz, H., Stembalska, A., Gil, J., . . . Sasiadek, M. M. (2011). Irrelevance of CHEK2 variants to diagnosis of breast/ovarian cancer predisposition in Polish cohort. *J Appl Genet*, 52(2), 185–191. <https://doi.org/10.1007/s13353-010-0013-1>
58. Nachef, M., Ali, A. K., Almutairi, S. M., & Lee, S. H. (2021). Targeting SLC1A5 and SLC3A2/SLC7A5 as a Potential Strategy to Strengthen Anti-Tumor Immunity in the Tumor Microenvironment. *Front Immunol*, 12, 624324. <https://doi.org/10.3389/fimmu.2021.624324>
59. Nakajima, H., Ishigaki, Y., Xia, Q. S., Ikeda, T., Yoshitake, Y., Yonekura, H., . . . Motoo, Y. (2010). Induction of HITS, a newly identified family with sequence similarity 107 protein (FAM107B), in cancer cells by heat shock stimulation. *Int J Oncol*, 37(3), 583–593. https://doi.org/10.3892/ijo_00000707
60. Nakajima, H., & Koizumi, K. (2014). Family with sequence similarity 107: A family of stress responsive small proteins with diverse functions in cancer and the nervous system (Review). *Biomed Rep*, 2(3), 321–325. <https://doi.org/10.3892/br.2014.243>
61. Nakajima, H., Koizumi, K., Tanaka, T., Ishigaki, Y., Yoshitake, Y., Yonekura, H., . . . Motoo, Y. (2012). Loss of HITS (FAM107B) expression in cancers of multiple organs: tissue microarray analysis. *Int J Oncol*, 41(4), 1347–1357. <https://doi.org/10.3892/ijo.2012.1550>
62. Narod, S. A. (2002). Modifiers of risk of hereditary breast and ovarian cancer. *Nat Rev Cancer*, 2(2), 113–123. <https://doi.org/10.1038/nrc726>
63. Narod, S. A., & Lynch, H. T. (2007). CHEK2 mutation and hereditary breast cancer. *J Clin Oncol*, 25(1), 6–7. <https://doi.org/10.1200/JCO.2006.08.8229>
64. Orliac, E. J., Trejo Banos, D., Ojavee, S. E., Läll, K., Mägi, R., Visscher, P. M., & Robinson, M. R. (2022). Improving GWAS discovery and genomic prediction accuracy in biobank data. *Proc Natl Acad Sci U S A*, 119(31), e2121279119. <https://doi.org/10.1073/pnas.2121279119>
65. Pankratov, V., Montinaro, F., Kushniarevich, A., Hudjashov, G., Jay, F., Saag, L., . . . Metspalu, M. (2020). Differences in local population history at the finest level: the case of the Estonian population. *Eur J Hum Genet*, 28(11), 1580–1591. <https://doi.org/10.1038/s41431-020-0699-4>
66. Pavlovica, K., Irmejs, A., Noukas, M., Palover, M., Kals, M., Tonisson, N., . . . Gardovskis, J. (2022). Spectrum and frequency of CHEK2 variants in breast cancer affected and general population in the Baltic states region, initial results and literature review. *Eur J Med Genet*, 65(5), 104477. <https://doi.org/10.1016/j.ejmg.2022.104477>

67. Phelan, C. M., Kuchenbaecker, K. B., Tyrer, J. P., Kar, S. P., Lawrenson, K., Winham, S. J., . . . Pharoah P. D. P. (2017). Identification of 12 new susceptibility loci for different histotypes of epithelial ovarian cancer. *Nat Genet*, *49*(5), 680–691. <https://doi.org/10.1038/ng.3826>
68. Plakhins, G., Irmejs, A., Gardovskis, A., Subatniece, S., Rozite, S., Bitina, M., . . . Gardovskis, J. (2011). Genotype-phenotype correlations among BRCA1 4153delA and 5382insC mutation carriers from Latvia. *BMC Med Genet*, *12*, 147. <https://doi.org/10.1186/1471-2350-12-147>
69. Plonis, J., Kalniete, D., Nakazawa-Miklasevica, M., Irmejs, A., Vjaters, E., Gardovskis, J., & Miklasevics, E. (2015). The. *Balkan J Med Genet*, *18*(2), 33–36. <https://doi.org/10.1515/bjmg-2015-0083>
70. Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*, *38*(8), 904–909. <https://doi.org/10.1038/ng1847>
71. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., . . . Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*, *81*(3), 559–575. <https://doi.org/10.1086/519795>
72. Rebbeck, T. R., Friebel, T. M., Friedman, E., Hamann, U., Huo, D., Kwong, A., . . . Nathanson, K. L. (2018). Mutational spectrum in a worldwide study of 29,700 families with BRCA1 or BRCA2 mutations. *Hum Mutat*, *39*(5), 593–620. <https://doi.org/10.1002/humu.23406>
73. Rebbeck, T. R., Mitra, N., Wan, F., Sinilnikova, O. M., Healey, S., McGuffog, L., . . . Andrulis, I. (2015). Association of type and location of BRCA1 and BRCA2 mutations with risk of breast and ovarian cancer. *JAMA*, *313*(13), 1347–1361. <https://doi.org/10.1001/jama.2014.5985>
74. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., . . . ACMG Laboratory Quality Assurance Committee (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*, *17*(5), 405–424. <https://doi.org/10.1038/gim.2015.30>
75. Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., & Müller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, *12*, 77. <https://doi.org/10.1186/1471-2105-12-77>
76. Rovite, V., Wolff-Sagi, Y., Zaharenko, L., Nikitina-Zake, L., Grens, E., & Klovinis, J. (2018). Genome Database of the Latvian Population (LGDB): Design, Goals, and Primary Results. *J Epidemiol*, *28*(8), 353–360. <https://doi.org/10.2188/jea.JE20170079>

77. Rowley, S. M., Mascarenhas, L., Devereux, L., Li, N., Amarasinghe, K. C., Zethoven, M., . . . Campbell, I. G. (2019). Population-based genetic testing of asymptomatic women for breast and ovarian cancer susceptibility. *Genet Med*, *21*(4), 913–922. <https://doi.org/10.1038/s41436-018-0277-0>
78. Sokolenko, A. P., Bogdanova, N., Kluzniak, W., Preobrazhenskaya, E. V., Kuligina, E. S., Iyevleva, A. G., . . . Imyanitov, E. N. (2014). Double heterozygotes among breast cancer patients analyzed for BRCA1, CHEK2, ATM, NBN/NBS1, and BLM germ-line mutations. *Breast Cancer Res Treat*, *145*(2), 553–562. <https://doi.org/10.1007/s10549-014-2971-1>
79. Su, X., Malouf, G. G., Chen, Y., Zhang, J., Yao, H., Valero, V., . . . Esteva, F. J. (2014). Comprehensive analysis of long non-coding RNAs in human breast cancer clinical subtypes. *Oncotarget*, *5*(20), 9864–9876. <https://doi.org/10.18632/oncotarget.2454>
80. Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*, *71*(3), 209–249. <https://doi.org/10.3322/caac.21660>
81. Tamboom, K., Kaasik, K., Aršavskaja, J., Tekkel, M., Lilleorg, A., Padrik, P., . . . Veidebaum, T. (2010). BRCA1 mutations in women with familial or early-onset breast cancer and BRCA2 mutations in familial cancer in Estonia. *Hered Cancer Clin Pract*, *8*(1), 4. <https://doi.org/10.1186/1897-4287-8-4>
82. Team, R. (2020). RStudio: Integrated Development Environment for R; RStudio, PBC: Boston, MA, USA, 2020; Available online: <http://www.rstudio.com/> (accessed on 27 May 2023).
83. Team, R. C. (2020). R: A Language and Environment for Statistical Computing; R Foundation for Statistical Computing: Vienna, Austria, 2020; Available online: <https://www.R-project.org/> (accessed on 27 May 2023).
84. Therneau. (2020). A Package for Survival Analysis in R. R package version 3.2–3. In. Available online: <https://CRAN.R-project.org/package=survival> (accessed on 17 October 2023).
85. Thulesius, H. O., Lindgren, A. C., Olsson, H. L., & Håkansson, A. (2004). Diagnosis and prognosis of breast and ovarian cancer – a population-based study of 234 women. *Acta Oncol*, *43*(2), 175–181. <https://doi.org/10.1080/02841860310022481>
86. Tikhomirova, L., Sinicka, O., Smite, D., Eglitis, J., Hodgson, S. V., & Stengrevics, A. (2005). High prevalence of two BRCA1 mutations, 4154delA and 5382insC, in Latvia. *Fam Cancer*, *4*(2), 77–84. <https://doi.org/10.1007/s10689-004-2758-3>
87. Turnbull, C., Seal, S., Renwick, A., Warren-Perry, M., Hughes, D., Elliott, A., . . . Rahman, N. (2012). Gene-gene interactions in breast cancer susceptibility. *Hum Mol Genet*, *21*(4), 958–962. <https://doi.org/10.1093/hmg/ddr525>

88. Van den Boom, J., Wolter, M., Blaschke, B., Knobbe, C. B., & Reifenberger, G. (2006). Identification of novel genes associated with astrocytoma progression using suppression subtractive hybridization and real-time reverse transcription-polymerase chain reaction. *Int J Cancer*, *119*(10), 2330–2338. <https://doi.org/10.1002/ijc.22108>
89. Van Geldermalsen, M., Wang, Q., Nagarajah, R., Marshall, A. D., Thoeng, A., Gao, D., . . . Holst, J. (2016). ASCT2/SLC1A5 controls glutamine uptake and tumour growth in triple-negative basal-like breast cancer. *Oncogene*, *35*(24), 3201–3208. <https://doi.org/10.1038/onc.2015.381>
90. Wang, L., Darling, J., Zhang, J. S., Liu, W., Qian, J., Bostwick, D., . . . Smith, D. I. (2000). Loss of expression of the DRR 1 gene at chromosomal segment 3p21.1 in renal cell carcinoma. *Genes Chromosomes Cancer*, *27*(1), 1–10. [https://doi.org/10.1002/\(sici\)1098-2264\(200001\)27:1<1::aid-gcc1>3.0.co;2-6](https://doi.org/10.1002/(sici)1098-2264(200001)27:1<1::aid-gcc1>3.0.co;2-6)
91. Wang, Y., Zhao, W., Liu, X., Guan, G., & Zhuang, M. (2019). ARL3 is downregulated and acts as a prognostic biomarker in glioma. *J Transl Med*, *17*(1), 210. <https://doi.org/10.1186/s12967-019-1914-3>
92. Watanabe, K., Taskesen, E., van Bochoven, A., & Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. *Nat Commun*, *8*(1), 1826. <https://doi.org/10.1038/s41467-017-01261-5>
93. Wiesmueller, F., Kopke, J., Aust, D., Roy, J., Dahl, A., Pilarsky, C., & Grützmann, R. (2019). Silenced ZNF154 Is Associated with Longer Survival in Resectable Pancreatic Cancer. *Int J Mol Sci*, *20*(21). <https://doi.org/10.3390/ijms20215437>
94. Yang, W., Zhang, T., Song, X., Dong, G., Xu, L., & Jiang, F. (2022). SNP-Target Genes Interaction Perturbing the Cancer Risk in the Post-GWAS. *Cancers (Basel)*, *14*(22). <https://doi.org/10.3390/cancers14225636>
95. Ye, M., Li, L., Liu, D., Wang, Q., Zhang, Y., & Zhang, J. (2021). Identification and validation of a novel zinc finger protein-related gene-based prognostic model for breast cancer. *PeerJ*, *9*, e12276. <https://doi.org/10.7717/peerj.12276>
96. Zhao, Z., Guo, Y., Liu, Y., Sun, L., Chen, B., Wang, C., . . . Li, X. (2021). Individualized lncRNA differential expression profile reveals heterogeneity of breast cancer. *Oncogene*, *40*(27), 4604–4614. <https://doi.org/10.1038/s41388-021-01883-6>

Pateicības

Pirmkārt, es vēlos izteikt sirsnīgu pateicību saviem promocijas darba vadītājiem **Edvīnam Miklaševičam** un **Baibai Vilnei** par atbalstu un vadību manā doktorantūras ceļā. Pateicos **E. Miklaševičam** par man piedāvāto iespēju izstrādāt promocijas darbu, par pētījuma tēmas ideju un par to, ka man iemācījāt patstāvību pētniecībā. **B. Vilnei** esmu parādā dziļu pateicību par ierašanos RSU un par to, ka pieņēmāt mani zem sava “spārna” un pacietīgi vadījāt cauri sarežģītajai bioinformātikas pasaulei. Varu droši apgalvot, ka bez Jūsu palīdzības nebūtu iespējams pabeigt analīzes un izstrādāt šo promocijas darbu.

Izsaku arī pateicību savai zinātniskajai konsultantei **Jeļenai Maksimenko** un kolēģiem no Paula Stradiņa Klīniskās universitātes slimnīcas Krūts ķirurģijas nodaļas par nenovērtējamo palīdzību pētījuma dalībnieku atlasē un paraugu vākšanā.

Sirsnīgu pateicību izsaku savai recenzentei **asoc. prof. Madarai Auzenbahai** par veltīto laiku, lai recenzētu manu promocijas darbu, kā arī sniegtajiem vērtīgajiem komentāriem un padomiem. Jūsu ieguldījums neapšaubāmi ir uzlabojis šī darba kvalitāti. Izsaku pateicību arī visiem promocijas darba apspriešanas komisijas locekļiem par darba izvērtēšanu un vērtīgajiem ieteikumiem.

Vēlos izteikt pateicību saviem kolēģiem no Rīgas Stradiņa universitātes Onkoloģijas un molekulārās ģenētikas institūta par nepārtraukto atbalstu un padomiem darba tapšanā.

Īpašu paldies saku saviem kolēģiem Bērnu klīniskajā universitātes slimnīcā par degsmi, ko mēs kopīgi izjūtam pret ģenētiku, par nebeidzamo atbalstu un motivāciju, kas ir bijis mana spēka avots grūtos brīžos.

Esmu arī ļoti pateicīga kolēģiem no Tartu universitātes Genomikas institūta par viņu sadarbības garu, nenovērtējamiem padomiem, apmācībām un dalīšanos ar pieredzi GWAS analīzē. Bez jūsu atbalsta es nevarētu pabeigt šo

darbu. Īpašu pateicību vēlos izteikt *Tiit Nikopensius, Reedik Mägi, Krista Fischer, Triin Laisk* un *Urmo Vösa*.

Esmu ļoti pateicīga saviem draugiem un ģimenei par nebeidzamo atbalstu un ticību maniem spēkiem. Īpašs paldies pienākas manam vīram par viņa apbrīnojamo pacietību, nenogurstošo sapratni un nepārtraukto iedrošinājumu visa doktorantūras ceļa garumā.

Visbeidzot, es izsaku pateicību visiem, kas mani atbalstīja un uzmundrināja ceļā uz doktora grādu. Katram no jums ir bijusi nozīmīga loma manas akadēmiskās un personīgās izaugsmes veicināšanā, un par to esmu jums patiesi pateicīga.