The Cardiovascular Impact and Genetics of Pericardial Adiposity

Supplementary Information

Table of Contents

Supplementary Methods

Fine-tuning of the deep learning model

All four-chamber images were resized to 168 x 208 pixels and normalized using mean and standard deviation estimates from ImageNet^{[1](https://paperpile.com/c/c7arBe/UeEk)}. The model was fine-tuned with a batch size of four. Augmentations with rotations up to 90 degrees were applied. Gradient clipping was used to limit the size of the gradients. The learning rate was chosen as $1x10⁻⁴$ based on the learning rate finder and the model was trained for 50 cycles while monitoring the Dice score. We applied the fine-tuned model to randomly selected four-chamber CMR images from all remaining 45,263 UKB participants after excluding participants in the training, validation and test sets.

Additional quality control for CMR images

A small proportion of the UK Biobank four-chamber CMR images were misaligned (not true four-chamber images) or of poor quality e.g. due to imaging artifacts. To exclude these images from final analyses, we first manually annotated the four cardiac chambers (left ventricle, left atrium, right ventricle, right atrium) in the same training data that was used for PAT model training. The deep learning model based on Resnet50 was fine-tuned with additional labels for the four cardiac chambers and used to infer the segmentation of the cardiac chambers in addition to PAT^{[2](https://paperpile.com/c/c7arBe/MFQW)}. We excluded any images in which any of the cardiac chambers was predicted to have an area less than $5cm²$.

FinnGen Ethics Statement

Patients and control subjects in FinnGen provided informed consent for biobank research, based on the Finnish Biobank Act. Alternatively, separate research cohorts, collected prior the Finnish Biobank Act came into effect (in September 2013) and start of FinnGen (August 2017), were collected based on study-specific consents and later transferred to the Finnish biobanks

after approval by Fimea (Finnish Medicines Agency), the National Supervisory Authority for Welfare and Health. Recruitment protocols followed the biobank protocols approved by Fimea. The Coordinating Ethics Committee of the Hospital District of Helsinki and Uusimaa (HUS) statement number for the FinnGen study is Nr HUS/990/2017.

The FinnGen study is approved by Finnish Institute for Health and Welfare (permit numbers: THL/2031/6.02.00/2017, THL/1101/5.05.00/2017, THL/341/6.02.00/2018, THL/2222/6.02.00/2018, THL/283/6.02.00/2019, THL/1721/5.05.00/2019 and THL/1524/5.05.00/2020), Digital and population data service agency (permit numbers: VRK43431/2017-3, VRK/6909/2018-3, VRK/4415/2019-3), the Social Insurance Institution (permit numbers: KELA 58/522/2017, KELA 131/522/2018, KELA 70/522/2019, KELA 98/522/2019, KELA 134/522/2019, KELA 138/522/2019, KELA 2/522/2020, KELA 16/522/2020), Findata permit numbers THL/2364/14.02/2020, THL/4055/14.06.00/2020, THL/3433/14.06.00/2020, THL/4432/14.06/2020, THL/5189/14.06/2020, THL/5894/14.06.00/2020, THL/6619/14.06.00/2020, THL/209/14.06.00/2021, THL/688/14.06.00/2021, THL/1284/14.06.00/2021, THL/1965/14.06.00/2021, THL/5546/14.02.00/2020, THL/2658/14.06.00/2021, THL/4235/14.06.00/2021, Statistics Finland (permit numbers: TK-53-1041-17 and TK/143/07.03.00/2020 (earlier TK-53-90-20) TK/1735/07.03.00/2021, TK/3112/07.03.00/2021) and Finnish Registry for Kidney Diseases permission/extract from the meeting minutes on 4th July 2019.

The Biobank Access Decisions for FinnGen samples and data utilized in FinnGen Data Freeze 10 include: THL Biobank BB2017_55, BB2017_111, BB2018_19, BB_2018_34, BB_2018_67, BB2018_71, BB2019_7, BB2019_8, BB2019_26, BB2020_1, BB2021_65, Finnish Red Cross Blood Service Biobank 7.12.2017, Helsinki Biobank HUS/359/2017, HUS/248/2020, HUS/150/2022 § 12, §13, §14, §15, §16, §17, §18, and §23, Auria Biobank AB17-5154 and

amendment #1 (August 17 2020) and amendments BB_2021-0140, BB_2021-0156 (August 26 2021, Feb 2 2022), BB_2021-0169, BB_2021-0179, BB_2021-0161, AB20-5926 and amendment #1 (April 23 2020)and it´s modification (Sep 22 2021), Biobank Borealis of Northern Finland 2017 1013, 2021 5010, 2021 5018, 2021 5015, 2021 5023, 2021 5017, 2022 6001, Biobank of Eastern Finland 1186/2018 and amendment 22 § /2020, 53§/2021, 13§/2022, 14§/2022, 15§/2022, Finnish Clinical Biobank Tampere MH0004 and amendments (21.02.2020 & 06.10.2020), §8/2021, §9/2022, §10/2022, §12/2022, §20/2022, §21/2022, §22/2022, §23/2022, Central Finland Biobank 1-2017, and Terveystalo Biobank STB 2018001 and amendment 25th Aug 2020, Finnish Hematological Registry and Clinical Biobank decision 18th June 2021, Arctic biobank P0844: ARC_2021_1001.

Endpoint definitions in FinnGen

We used the following FinnGen endpoints from Data Freeze 11: E4_DM2 (Type 2 diabetes), I9_AF (atrial fibrillation or flutter), I9_HEARTFAIL (heart failure), I9_CHD (Coronary artery disease), and I9_STR (stroke). FinnGen endpoint definitions incorporate International Classification of Diseases (ICD) codes from hospital, outpatient and cause-of-death registries; ICD or Social Insurance Institution (KELA) classification based medication reimbursement codes; and Anatomical Therapeutic Chemical codes for medication purchases. The definitions are publicly available at <https://www.finngen.fi/en/researchers/clinical-endpoints> and on <https://risteys.finregistry.fi>.

Gene prioritization

We used two approaches to prioritize potentially causative genes in the association loci: nearest protein-coding gene and Polygenic Priority Score (PoPS) [3](https://paperpile.com/c/c7arBe/Qt3I) . To compute gene-specific PoPS, we first performed a gene-based analysis with MAGMA v1.10 using the default snp-wise=mean gene analysis model^{[4](https://paperpile.com/c/c7arBe/wS6y)}. The raw gene-based results from MAGMA and the full set of default features provided by the PoPS authors were then used to calculate gene-specific PoPS. The

gene with the highest PoPS within +/- 500kb of the lead variant was designated as the gene prioritized by PoPS.

Genome-wide association study meta-analysis

We performed a p-value based sample size weighted meta-analysis of the UKB GWAS and a previously published GWAS meta-analysis (n = 12,204)^{[5,6](https://paperpile.com/c/c7arBe/NAuT+Suil)} using METAL (version release 2020-05-05) without genomic control. We evaluated for the presence of inflation based on the LD score regression intercept for the novel meta-analysis.

Polygenic score analyses

We constructed a polygenic score (PGS) for PAT using the summary statistics from the UKB GWAS and the PRS-CS-auto tool with a UK Biobank European ancestry LD reference panel and HapMap3 variants^{[7](https://paperpile.com/c/c7arBe/WlT0)}. We calculated individual-level PGSs using the resulting weights in Finn Gen study participants with PLINK $v2.0^8$ $v2.0^8$. We evaluated the associations of the PAT PGS with cardiovascular diseases registered at any point before or after DNA sampling as outcomes, using logistic regression models with sex, age at the end of study follow-up or death, genomic principal components 1–5 and the genotyping array as basic covariates.

Supplementary Results

Genomic inflation and LD Score Regression Intercept

The genomic control inflation statistic (lambda) of the genome-wide association study of PAT in UKB was modestly elevated at 1.096 (Supplementary Figure 7), but the linkage disequilibrium score regression intercept was 1.00, consistent with true polygenic signal. Similarly, in the meta-analysis of the UKB GWAS and a previously published PAT GWAS meta-analysis^{[6](https://paperpile.com/c/c7arBe/Suil)}, the genomic inflation factor was 1.103 and the LD score regression intercept was 1.0105 (SE 0.0075).

Supplementary Figures

Supplementary Figure 1: Study flowchart

Supplementary Figure 2: The associations of pericardial adipose tissue with age, sex, and anthropometric measurements

A. The association of PAT with age during the imaging visit. B. The distribution of PAT in women (red) and men (blue). C. The correlations of anthropometric measurements with pericardial adipose tissue. Pearson's correlation coefficient is shown for all pairs of continuous traits. Abbreviations: BMI = body mass index, PAT = pericardial adipose tissue, WHR = waist to hip ratio.

Supplementary Figure 3: The associations of pericardial adipose tissue with incident diseases

A. Adjusted for basic covariates

B. Adjusted for basic covariates and BMI

The associations of PAT with incident diseases were tested using Cox proportional hazards models with time from imaging to diagnosis or censoring as the outcome and PAT (standard deviation scaled), sex, and age as the predictors. Body mass index (BMI) was included as an additional covariate in BMI-adjusted models. Participants with the corresponding prevalent disease at the time of imaging were excluded from analyses. Effect estimates are reported for PAT. CI = confidence interval.

Supplementary Figure 4: The associations of percentile-stratified pericardial adipose tissue with incident diseases

0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 Hazard ratio in individuals with PAT > 90% percentile

A. Adjusted for basic covariates

Stroke

The associations of PAT with incident diseases were tested using Cox proportional hazards models with time from imaging to diagnosis or censoring as the outcome and PAT percentile (over 90th percentile vs. others), sex, and age as the predictors. Participants with the corresponding prevalent disease at the time of imaging were excluded from analyses. Body mass index (BMI) was included as an additional covariate in BMI-adjusted models. Hazard

 $1.09(0.72 - 1.65)$

0.68

ratios estimates are reported for the PAT strata (> 90th percentile compared with PAT ≤ 90th percentile)

Supplementary Figure 5: Cumulative incidences of cardiovascular diseases stratified by PAT decile at baseline

Atrial fibrillation or flutter

Coronary artery disease

Type 2 diabetes

Supplementary Figure 6: Locus plots from the genome-wide association study of pericardial adipose tissue in UK Biobank

Supplementary Figure 7: Quantile-quantile plot of p-values from the genome-wide association study of pericardial adipose tissue in UK Biobank

A genome-wide association study of pericardial adipose tissue was performed in 41,494 UK Biobank participants. Expected and observed p-values are shown on the -log10 scale.

Supplementary Figure 8: Manhattan plot of the meta-analysis of genome-wide association studies of pericardial adipose tissue

A p-value based sample size weighted meta-analysis of the UK Biobank genome-wide association study and a previously published GWAS meta-analysis by Chu et al (n = 12,204) was performed using METAL, including up to 53,698 total participants. Each variant is plotted as a data point, with the corresponding -log₁₀(P) shown on the y-axis and the genomic position shown on the x-axis grouped by chromosomes. The genome-wide significance threshold ($P =$ $5*10⁻⁸$) is shown with a darker dashed line and a suggestive threshold (P = 1*10⁻⁵) is shown with a lighter dashed line. Genomic loci with at least one variant reaching genome-wide significance are labeled with the name of the nearest gene, and all variants within 500 kilobases of the lead variant are coloured in a darker blue for visualization purposes. The y-axis is truncated to only show variants with a P-value ≤ 0.1 .

Supplementary Figure 9: Quantile-quantile plot of the meta-analysis of genome-wide association studies of pericardial adipose tissue

A p-value based sample size weighted meta-analysis of the UK Biobank genome-wide association study and a previously published GWAS meta-analysis by Chu et al (n = 12,204) was performed using METAL, including up to 53,698 total participants. Expected and observed p-values are shown on the -log10 scale.

Supplementary Tables

Supplementary Table 1: Characteristics of the study sample at the time of

imaging and stratified by self-reported ethnic background

Baseline characteristics at the time of imaging are shown for UK Biobank cardiac magnetic resonance imaging substudy participants with automated pericardial adipose tissue (PAT) area quantitation from four-chamber images. Self-reported ancestry was used to stratify participants. Anthropometric measurements were taken during the imaging visit. Prevalent diseases were ascertained based on a combination of self-reported data and International Classification of Diseases codes.

Supplementary Table 2: UK Biobank Disease Definitions

Cardiovascular diseases in UK Biobank were defined using a combination of International Classification of Diseases (ICD) codes, self-report, and procedure codes. AF = atrial fibrillation or flutter, CAD = coronary artery disease, HF = heart failure, T2D = type 2 diabetes.

Supplementary Table 3: Prevalent disease association results in the UK Biobank

In linear regression models, standard deviation scaled pericardial adipose tissue was used as the outcome and prevalent disease status (at the time of imaging), age and sex were used as the predictors. The effect (beta) estimates correspond to the differences in pericardial adipose tissue (in SD units) between individuals with prevalent disease and without prevalent disease.

Supplementary Table 4: The associations of pericardial adipose tissue with incident diseases

The associations of PAT with incident diseases were tested using Cox proportional hazards models with time from imaging to diagnosis or censoring as the outcome and PAT (standard deviation scaled), sex, and age as the predictors. Body mass index (BMI) was included as an additional covariate in BMI-adjusted models. Participants with the corresponding prevalent disease at the time of imaging were excluded from analyses. Effect estimates are reported for PAT. CI = confidence interval, HR = hazard ratio.

Supplementary Table 5: The associations of percentile-stratified pericardial adipose tissue with incident diseases

The associations of PAT with incident diseases were tested using Cox proportional hazards models with time from imaging to diagnosis or censoring as the outcome and PAT percentile (over 90th percentile vs. others), sex, and age as the predictors. Participants with the corresponding prevalent disease at the time of imaging were excluded from analyses. Body mass index (BMI) was included as an additional covariate in BMI-adjusted models. Effect estimates are reported for the PAT strata (> 90th percentile compared with PAT ≤ 90th percentile). CI = confidence interval, HR = hazard ratio.

Supplementary Table 6: Polygenic priority scores for genes in genome-wide significant loci in the genome-wide association study of pericardial adipose tissue in UK Biobank

Polygenic Priority Scores (PoPS) were calculated for all genes within +/- 500,000kb of the lead variant in each genome-wide significant locus. "Min pos" and "Max pos" correspond to the locus boundaries. Start and End correspond to the regions of each gene. TSS = Transcription Start Site.

Supplementary Table 7: Genetic correlations between pericardial adipose tissue and anthropometric traits

Genetic correlations between pericardial adipose tissue (PAT) and anthropometric measurements were calculated based on the summary statistics from genome-wide association studies in the UKB imaging substudy, using LD Score Regression with HapMap3 variants and a European ancestry reference panel. rg = genetic correlation, se = standard error.

Supplementary Table 8: Lead variants from the meta-analysis of genome-wide association studies of pericardial adipose tissue

A p-value based sample size weighted meta-analysis of the UK Biobank genome-wide association study and a previously published GWAS meta-analysis by Chu et al (n = 12,204) was performed using METAL, including up to 53,698 total participants. Lead variants, nearest genes and variant functions are shown. Variant allele frequencies (AF) are based on UK Biobank data only. CHROM = chromosome, POS = position in the GRCh37 build.

Supplementary Table 9: FinnGen participant characteristics

Sample characteristics are shown for 453,733 individuals from data freeze 11. Cardiovascular diseases were defined using a combination of International Classification of Diseases (ICD) codes from specialist inpatient, outpatient and cause-of-death registries, procedure codes, and medication reimbursement codes. Sex, age at genotyping, and body mass index (BMI) were ascertained at study recruitment. Disease-specific outcomes include both prevalent and incident diseases before the end of the follow-up.

Supplementary Table 10: The predictive utility of a polygenic score for pericardial adipose tissue in FinnGen

A polygenic score (PGS) for pericardial adipose tissue (PAT) was constructed using PRS-CS based on summary statistics from the genome-wide association study of PAT in UK Biobank and subsequently applied to 453,733 participants in the FinnGen study. The associations of the PAT PGS with cardiovascular diseases were evaluated using logistic regression models with sex, age at the end of study follow-up or death, genomic principal components 1–5 and the genotyping array as basic covariates. In body mass index (BMI) adjusted models, BMI was included as an additional covariate. Effect estimates are shown per SD increment in PAT PGS. CI = confidence interval, OR = odds ratio.

Supplementary References

- 1. Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L. ImageNet: A large-scale [hierarchical](http://paperpile.com/b/c7arBe/UeEk) image database. In: 2009 IEEE Conference on Computer Vision and Pattern [Recognition.](http://paperpile.com/b/c7arBe/UeEk) 2009. p. [248–255.](http://paperpile.com/b/c7arBe/UeEk)
- 2. He K, Zhang X, Ren S, Sun J. Deep residual learning for image [recognition.](http://paperpile.com/b/c7arBe/MFQW) In: [Proceedings](http://paperpile.com/b/c7arBe/MFQW) of the IEEE conference on computer vision and pattern recognition. 2016. p. [770–778.](http://paperpile.com/b/c7arBe/MFQW)
- 3. Weeks EM, Ulirsch JC, Cheng NY, Trippe BL, Fine RS. Leveraging polygenic [enrichments](http://paperpile.com/b/c7arBe/Qt3I) of gene features to predict genes [underlying](http://paperpile.com/b/c7arBe/Qt3I) complex traits and diseases. *MedRxiv* [Internet]. 2020; Available from: <https://www.medrxiv.org/content/10.1101/2020.09.08.20190561.abstract>
- 4. de Leeuw CA, Mooij JM, Heskes T. MAGMA: [generalized](http://paperpile.com/b/c7arBe/wS6y) gene-set analysis of GWAS data. *PLoS Comput Biol* [Internet]. [2015;Available](http://paperpile.com/b/c7arBe/wS6y) from: <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004219>
- 5. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient [meta-analysis](http://paperpile.com/b/c7arBe/NAuT) of genomewide association scans. *Bioinformatics*. [2010;26:2190–2191.](http://paperpile.com/b/c7arBe/NAuT)
- 6. Chu AY, Deng X, Fisher VA, Drong A, Zhang Y, [Feitosa](http://paperpile.com/b/c7arBe/Suil) MF, Liu C-T, Weeks O, Choh AC, Duan Q, Dyer TD, Eicher JD, Guo X, [Heard-Costa](http://paperpile.com/b/c7arBe/Suil) NL, Kacprowski T, Kent JW Jr, Lange LA, Liu X, Lohman K, Lu L, Mahajan A, [O'Connell](http://paperpile.com/b/c7arBe/Suil) JR, Parihar A, Peralta JM, Smith AV, Zhang Y, Homuth G, [Kissebah](http://paperpile.com/b/c7arBe/Suil) AH, Kullberg J, Laqua R, Launer LJ, Nauck M, Olivier M, Peyser PA, Terry JG, [Wojczynski](http://paperpile.com/b/c7arBe/Suil) MK, Yao J, Bielak LF, Blangero J, Borecki IB, Bowden DW, Carr JJ, [Czerwinski](http://paperpile.com/b/c7arBe/Suil) SA, Ding J, Friedrich N, Gudnason V, Harris TB, Ingelsson E, Johnson AD, Kardia SLR, [Langefeld](http://paperpile.com/b/c7arBe/Suil) CD, Lind L, Liu Y, Mitchell BD, Morris AP, Mosley TH Jr, Rotter JI, Shuldiner AR, Towne B, Völzke H, [Wallaschofski](http://paperpile.com/b/c7arBe/Suil) H, Wilson JG, Allison M, Lindgren CM, Goessling W, Cupples LA, [Steinhauser](http://paperpile.com/b/c7arBe/Suil) ML, Fox CS. Multiethnic genome-wide [meta-analysis](http://paperpile.com/b/c7arBe/Suil) of ectopic fat depots identifies loci associated with adipocyte development and differentiation. *Nat Genet*. [2017;49:125–130.](http://paperpile.com/b/c7arBe/Suil)
- 7. Ge T, Chen C-Y, Ni Y, Feng Y-CA, Smoller JW. [Polygenic](http://paperpile.com/b/c7arBe/WlT0) prediction via Bayesian regression and continuous shrinkage priors. *Nat Commun*. [2019;10:1776.](http://paperpile.com/b/c7arBe/WlT0)
- 8. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. [Second-generation](http://paperpile.com/b/c7arBe/ZiiA) PLINK: rising to the challenge of larger and richer datasets. *[Gigascience](http://paperpile.com/b/c7arBe/ZiiA)*. 2015;4:7.
- 9. [Rosenbloom](http://paperpile.com/b/c7arBe/NQ4m) KR, Armstrong J, Barber GP, Casper J, Clawson H, Diekhans M, Dreszer TR, Fujita PA, [Guruvadoo](http://paperpile.com/b/c7arBe/NQ4m) L, Haeussler M, Harte RA, Heitner S, Hickey G, Hinrichs AS, Hubley R, [Karolchik](http://paperpile.com/b/c7arBe/NQ4m) D, Learned K, Lee BT, Li CH, Miga KH, Nguyen N, Paten B, Raney BJ, Smit AFA, Speir ML, Zweig AS, [Haussler](http://paperpile.com/b/c7arBe/NQ4m) D, Kuhn RM, Kent WJ. The UCSC Genome Browser database: 2015 update. *Nucleic Acids Res*. [2015;43:D670–81.](http://paperpile.com/b/c7arBe/NQ4m)
- 10. R Core Team. R: A language and [environment](http://paperpile.com/b/c7arBe/5Pwu) for statistical computing. R Foundation for Statistical [Computing,](http://paperpile.com/b/c7arBe/5Pwu) Vienna, Austria [Internet]. 2016. Available from: <https://cir.nii.ac.jp/crid/1574231874043578752>