Genetic Insights of Image-Based Traits: Analysis Pipeline for AI-based Phenotyping, Combined-GWAS, and Federated Learning with Application to the Human Face

XIANJING LIU[1,2,*], ZIYI XIONG[3,4,*,$], FAN LIU[3], TAMAR NIJSTEN[5], EPPO B. WOLVIUS[2], MANFRED KAYSER [3,#,$,], GENNADY V. ROSHCHUPKIN[1,4,#]

[1] Department of Radiology and Nuclear Medicine, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

[2] Department of Oral & Maxillofacial Surgery, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

[3] Department of Genetic Identification, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

[4] Department of Epidemiology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

[5] Department of Dermatology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

* These authors contributed equally to this work.

[#] These authors contributed equally to this work.

[$] Current affiliation: Department of Pathology and Clinical Bioinformatics, Erasmus MC University Medical Center Rotterdam, Rotterdam, the Netherlands

Corresponding author: Gennady Roshchupkin: g.roshchupkin@erasmusmc.nl

Keywords:

Biomedical images, Federated learning; AI phenotyping; C-GWAS; Human face

Funding sources:

The Rotterdam Study is funded through unrestricted research grants from Erasmus MC University Medical Center Rotterdam, Erasmus University Rotterdam, the Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The Generation R Study is conducted by the Erasmus MC in close collaboration with the School of Law and Faculty of Social Sciences of the Erasmus University Rotterdam, the Municipal Health Service Rotterdam area, Rotterdam, the Rotterdam Homecare Foundation, Rotterdam, and the Stichting Trombosedienst & Artsenlaboratorium Rijnmond (STAR-MDC), Rotterdam.

Conflicts of Interest:

The authors declare not to have a conflict of interest with this work.

IRB approval status: The Rotterdam Study has been approved by the Erasmus MC Medical Ethical Committee (MEC-02-1015), and by the Dutch Ministry of Health, Welfare and Sport (Population Screening Act, reference 3295110-1021635-PG). The Generation R Study was approved by the Erasmus MC Medical Ethics Committee (MEC 198.782/2001/31).

Abstract:

Image-derived phenotypes - traits extracted from biological images - capture rich morphological information and understanding their genetic basis is crucial for elucidating developmental mechanisms and linking genetic variation to complex visual traits, relevant in many areas of biomedical, evolutionary, and forensic research and applications. However, there are key limitations in the current methodology, such as in the degree of which the large image complexity is captured with the phenotyping methods and how the genetic analysis methods deal with the underlying large genetic complexity. Moreover, needed multi-cohort studies are constrained by privacy regulations often prohibit sharing individual image data across institutions. Here, we present a robust, scalable, privacy-preserving analysis pipeline for unveiling the genetic basis of image-based complex traits, integrating (i) AI-based phenotyping for automatically extracting large numbers of endophenotypes; (ii) Combined-GWAS (C-GWAS) for identifying genetic variants underlying the numerous endophenotypes; (iii) federated learning for training AI-based phenotyping models across multiple cohorts without sharing individual images; and (iv) explainable AI for image-based visualization of the identified genetic effects. In the first application, we analysed digital 3D facial images and genomic data from two European cohorts (N=7,309), extracted 200 image-derived facial endophenotypes, identified 43 significantly face associated genetic loci, including 12 novel ones, and replicated 70% of them in an independent European dataset (N=8,246). AI-based visualization of the identified genetic effects shows the involvement of many of these genetic loci in different parts of the face. Our study provides a generalizable, privacy-aware analysis framework for investigating the genetic basis of image-based complex traits implemented in a computationally efficient python package; its first application yielded new insights into the genetic architecture of facial shape variation.

Introduction

High-throughput biomedical imaging—ranging from 3D scans of bodies, body parts or organs to histopathology slides—captures complex structural and textural patterns that reflect underlying biology and disease. Yet, defining phenotypes that capture the full spectrum of a biological image is challenging with existing methodology. Several strategies have been proposed to tackle the challenges when it comes to unveiling the genetic basis of image-based complex traits, such as utilizing using each pixel or voxel within an image e.g. brain MRI[11] or prior-defined measurements e.g. as landmarks from facial images[9] or the region-of-interest volume from brain MRI scans[10] as phenotypes for genome-wide association studies (GWASs). Although these strategies yield commendable outcomes when applied to large sample size[9,10,11,12], the real potential lies in seeking solutions that uncover more comprehensive, data-driven image-based phenotypes that move beyond a priori definitions, and better reflect the underlying biological variation.

Recent breakthroughs in artificial intelligence (AI) models have demonstrated promising data-driven capabilities when analysing images by reducing high-dimensional images into low-dimensional representations of endophenotypes, while preserving all essential image information. This approach has proven fruitful in various contexts, such as unsupervised autoencoder-based endophenotypes from 3D brain images[14], 3D cardiac mesh images[15], 3D facial mesh images[80], self-supervised contrastive learning[16], and transfer learning[75] for 2D retinal fundus images. Until now, all GWASs of these AI-driven endophenotypes have applied the classical single-trait approach—testing each AI-driven endophenotype separately and in isolation from another. Yet across image modalities, these learned endophenotypes are inherently correlated, as result of underlying biological correlations, which requiring analysis methods that allow to detect genetic correlations among them, which single-trait GWAS cannot. To effectively combine single trait GWASs of

(un)correlated traits, some of us recently introduced a combined genome-wide association study approach named Combined-GWAS (C-GWAS), and demonstrated significantly improved power of C-GWAS relative to meta-analysis of single-traits GWASs in both, simulated data as well as real data applications such as to the human ear[76] and the human face[6]. Yet, C-GWAS has not been applied to image-based AI-generated endophenotypes.

An additional challenge is the standardization of AI-derived phenotypes across cohorts. Genetic studies of complex traits require very large sample sizes, often necessitating the combination of data from multiple institutions. For traditional traits, cohorts can share GWAS summary statistics derived from standardized measurements without exchanging raw image data. However, AI-derived phenotypes depend on cohort-specific training data. Consequently, models trained separately produce features that are not directly comparable. In addition, privacy regulations often restrict the sharing of raw imaging data needed to train a unified model. Federated learning addresses this issue by enabling all sites to collaboratively train a single AI model—without ever sharing individual image data.

Here, we present a robust, scalable, privacy-preserving analysis pipeline for unveiling the genetic basis of image-based complex traits that integrates (i) AI-based phenotyping for automatically extracting large numbers of endophenotypes from a biological image; (ii) Combined-GWAS (C-GWAS) for identifying genetic variants underlying the numerous endophenotypes; (iii) federated learning for training AI-based phenotyping models across multiple cohorts without sharing individual images; and (iv) explainable AI for image-based visualization of the identified genetic effects. To empirically exemplify its practical suitability, we applied our pipeline to the case example of the human face by using 3D digital facial images and genome-wide data of thousands of European individuals from several cohorts.

Methods and materials:

Study population

Rotterdam Study

The Rotterdam Study (RS) is a population-based cohort study of 14,926 participants aged 45 years and older, living in the same suburb of Rotterdam, the Netherlands[1,17]. The present study includes 3,995 participants (mean age=69.6 years, SD=9.5 years) of Dutch European ancestry, for whom high-resolution 3-dimensional digital face photographs were taken. The raw 3D facial images of participants were acquired using a 3D photographic scanning system manufactured by 3dMD (Atlanta (GA), USA). Participants were asked to keep their mouths closed and adopt a neutral expression during the acquisition of the 3D scans. Microarray-based genotyping was carried out according to the manufacturer's instructions with the Infinium II HumanHap 550K Genotyping BeadChip version 3 (Illumina, San Diego, California USA) at the Human Genomics Facility (HuGe-F) of the Department of Internal Medicine of Erasmus MC (www.glimdna.org)[2]. All SNPs were imputed using MACH software (www.sph.umich.edu/csg/abecasis/MaCH/) based on the 1000-Genomes Project reference population information[3]. After all quality controls, the current study included a total of 6,728,663 autosomal SNPs (MAF > 0.01, imputation R2 > 0.8, SNP call rate > 0.97, HWE > 0.0001). The Rotterdam Study has been approved by the Erasmus MC Medical Ethics Committee (registration number MEC 02.1015) and by the Dutch Ministry of Health, Welfare and Sport (Population Screening Act WBO, license number 1071272–159521 PG). The Rotterdam Study has been entered into the Netherlands National Trial Register (NTR; www.trialregister.nl) and into the WHO International Clinical Trials Registry Platform (ICTRP; www.who.int/ictrp/network/primary/en/) and under shared catalogue number

NTR6831. All participants provided written informed consent to participate in the study and to have their information obtained from treating physicians.

Generation R Study

The Generation R Study (GenR) is an ongoing population-based cohort study of pregnant women and their children from fetal life onwards[18]. The goal is to identify early environmental and genetic causes leading to normal and abnormal growth, development, and health. All women living in the study area of Rotterdam, the Netherlands, who delivered between April 2002 and January 2006 were eligible. The present study includes 3,314 participants (mean age=10.4 years, SD=1.5 years), for whom high-resolution 3-dimensional face digital photographs were taken using the 3dMD system. The participants were of multi-ethnic ancestry[35,36], and the largest ethnic groups were the Dutch-European (N=2,026, 61.1%), non-Dutch-European (N=243, 7.3%), Turkish (N=186, 5.6%) and Moroccan (N=175, 5.3%) groups.–Microarray-based genotyping[19] was performed at the Human Genomics Facility (HuGe-F) of Erasmus MC. All SNPs were imputed using MACH (version 1.0.15) software based on the 1000-Genomes Project reference population information. After all quality controls, the current study included a total of 7,416,225 autosomal SNPs (MAF > 0.01, imputation R2 > 0.8, SNP call rate > 0.98, HWE > 1.10e–6). The Generation R Study was approved by the Erasmus MC Medical Ethics Committee (MEC) (MEC 198.782/2001/31), and written informed consent was obtained from all participants themselves, or on that of their guardians/parents.

Facial endophenotypes from the AI-based modelling

The raw 3D facial image data were built into a template-based data set following a registration pipeline[20]. Then, we exploited a 3D graph autoencoder, a type of artificial neural networks to learn compressed representations of the input data in a self-supervised manner, to

perform dimensionality reduction on the facial images[21]. The autoencoder consists of an encoder and a decoder that can perform feature mapping in a non-linear manner. The encoder compresses the high-dimensional facial image into low-dimensional representations, while the decoder reconstructs the facial image from the representations. By minimizing the error between input and reconstructed facial images, the main facial morphology is captured in the low-dimensional representations. Implementation details of the autoencoder can be found elsewhere[20]. The dimension of representations was estimated to be 200, as a trade-off between the image reconstruction quality and the dimensional complexity[20]. We defined these 200 learned representations as facial endophenotypes, which were used in the GWAS as phenotypes.

Alignment of multi-cohort facial endophenotype via federated learning

Federated learning (FL)[65] (also known as collaborative learning) is a machine learning technique that trains a model via multiple independent sessions, each using its own dataset. This approach stands in contrast to traditional centralized training (CT) techniques where local datasets are merged into one training session[22]. A server was established upon clients of multiple (M) institutes, with a randomly initialized global AI model to be trained. In each round of the training, the global model is updated via following steps: 1) The server sends the global model to client #1, and then client #1 updates the received model with its own local data; 2) Client #1 sends the updated model back to the server; 3) repeat 1) and 2) for the rest (M-1) clients; 4) repeat 1), 2) and 3) for a prefixed number of rounds of the training. After the training is finished, a final global model was sent to clients. With these settings, we are able to train the proposed AI model simultaneous among multiple cohorts without sharing individual-level image data. Most importantly, since the AI model is optimised with facial image data from multiple institutes, the representing facial morphology of each

endophenotype is aligned across cohorts / institutes (Figure 1). This allows meta-analysis on traditional GWAS summary statistics of cohort data across institutes as well as C-GWAS.

To better understand the efficacy of the FL training strategy, we compared the performance of the FL and the CT approaches in terms of the error (Equation 1) between input and reconstructed facial mesh according to:

$$\text{Error} = \frac{\sum_{i}^{N_{vertex}} \sqrt[2]{(x_i - x_i')^2 + (y_i - y_i')^2 + (z_i - z_i')^2}}{N_{vertex}} \quad \text{(mm)} \qquad (1),$$

where N_{vertex} is the number of vertices of the facial mesh, (x_i, y_i, z_i) is a vertex of the input facial mesh while (x_i', y_i', z_i') is the corresponding vertex in the reconstructed facial mesh.

Single-trait GWAS, meta-analyses, and C-GWAS

In RS, traditional single-trait GWASs were conducted separately for the 200 standardised AI-based facial endophenotypes using 6,728,663 SNPs. GWAS was performed using linear regression under the additive genetic model while adjusting for sex, age, and for ancestry using the first four genomic PCs using Plink[4]. In GenR, GWASs were conducted separately for the 200 standardised AI-based facial endophenotypes and 7,416,225 SNPs using HASE[23] while adjusting for sex, age, and the first twenty genomic PCs. We included more PCs in GenR because it is a multi-ethnic cohort study. Inverse variance fixed-effect meta-analysis were carried out to combine GWAS results of RS and GenR for each of the 200 facial endophenotype, which included 7,733,921 autosomal SNPs with physical positions aligned according to 1NCBI human genome sequence build GRCh37.

Based on the single-trait GWAS meta-analysis outcomes, we performed C-GWAS. The C-GWAS R library is freely available at https://github.com/Fun-Gene/CGWAS. Details of the C-GWAS method have been described previously[6]. In brief, the null hypothesis (H0) for a tested SNP is the absence of any allelic effect on all traits, and the alternative hypothesis (H1) is that its allelic effects deviate from 0 for at least one of the multiple traits tested. C-

GWAS incorporates two different tests that originated from either the effect-based inversed covariance weighting, or the truncated Wald test to maximise statistical power. All resultant P-values from C-GWAS are adjusted using the getCoef function implemented in C-GWAS, which performs simulations (n simulations = 1e8) to guarantee that the null distribution of C-GWAS follows the uniform distribution in all quantiles. A common approach for handling correlated traits in GWAS involves performing separate GWASs for each single trait and selecting the minimum p-value (unadjusted MinGWAS). However, unadjusted MinGWAS is inflated under the null hypothesis (no effect on any trait), requiring stricter significance thresholds or adjustments for multiple correlated traits. In this study, we adjusted our unadjusted MinGWAS, i.e., minimal values of meta-analysis of 200 endophenotype GWASs, using the getCoef function in C-GWAS. This function employs simulation analyses to ensure that both C-GWAS and the adjusted MinGWAS follow a uniform distribution under the null. Consequently, our MinGWAS refers to the getCoef-calibrated minimum p-value, allowing direct comparison between the outcomes of C-GWAS, MinGWAS, and any single-trait GWAS. Thus, traditional genome-wide significance threshold of 5e-8 corresponds to our study-wide significance threshold for both C-GWAS and MinGWAS. Here, we used the default parameters of C-GWAS and applied C-GWAS to the summary statistics of the 200 AI-derived facial endophenotype GWASs in the meta-analysis of GenR and RS. Moreover, although not part of the core process of our proposed pipeline, we conducted additional C-GWAS separately in GenR and RS, to further illustrate how meta-analysis with our pipeline can improve the power.

Replication of C-GWAS outcomes

Aiming to replicate the C-GWAS derived top SNPs, we reached out and obtained the GWAS summary statistics (GWAS Catalog: GCP000044)[5] from 8,246 individuals of European

descent from a combination of three independently collected datasets from the United States (US; N = 4,680) and the United Kingdom (UK; N= 3,566). The data contains 126 sets of GWAS summary statistics corresponding to the two sets of 63 facial segments of US-driven or UK-driven meta-analyses respectively. These facial segments are derived as linear combinations of facial vectors within distinct regions, and they do not directly correspond to our AI-derived phenotypes based on whole facial vectors. Due to the differences in phenotyping methods between the discovery and replication studies, we used minimal p-values and C-GWAS embedded function getCoef to combine the evidence of replication from 126 GWASs and account for the correlation between facial segments. In brief, we first integrate 126 GWASs to one GWAS (MinGWAS) by taking the minimal p-values for each SNP. This is to test the null hypothesis that the SNP effect on all 126 facial segments are 0. Then we calibrate MinGWAS using getCoef function in C-GWAS, which guarantees the null of calibrated MinGWAS is the same as uniform 0-1 distribution and any of conventional GWAS[6]. We selected calibrated MinGWAS p-values of 43 lead SNPs as replicated evidence. $p<0.05$ indicates that the replicated SNP has effect on at least one facial segment with study-wide nominal significance.

Comparison of C-GWAS from AI-derived and landmark-based facial phenotypes

To directly compare the two phenotyping approaches in their genetic outcomes, we conducted single-trait GWAS separately for AI-derived phenotypes and landmark-distance-based phenotypes, using the same set of 3,761 RS samples with access to both phenotype sets. The landmarks distance phenotypes consist of 78 traits derived from 13 facial landmarks, as detailed in our previous study[9]. After conducting single-trait GWAS, we subsequently applied C-GWAS to the summary statistics of the 200 AI-derived phenotypes, and the 78 landmark-based phenotypes and compared the outcomes.

Visualizing genetic association effects in the human face

For interpreting the achieved genetic associations regarding their effect on the human face, we performed an explainable AI-based facial visualization for the C-GWAS-derived top SNPs. After the training of the autoencoder, each input facial mesh is represented by a latent datapoint (blue points in Figure 3) in the endophenotypes space. The endophenotype space can be considered as a vector subspace of $\square^{n}$ (n dimensions of the endophenotypes) with basis vectors $\overrightarrow{e_j}$ (j=1,2...n), in which any vector can be represented by $\vec{v} = w_1\overrightarrow{e_1}+w_2\overrightarrow{e_2}+ \ldots +w_n\overrightarrow{e_n}$ as a linear combination of basis vectors using their weightings w_j (j=1,2...n)[39,73] (Figure 3). By decoding a selected endophenotype (e.g., e_1) while keeping the rest unchanged, the decoder can perform sampling along an axis (e.g., $\overrightarrow{e_1}$) in the endophenotype space, and to reconstruct a series of facial shapes showing deformation related to the change of the selected endophenotype (Figure 3). Particularly, by selectively decoding a subset of endophenotypes with weightings w_1, w_2, …, w_n, the decoder can perform sampling along vector $\vec{v}$, and to reconstruct a series of facial shapes showing facial deformation related to the linear combination of multiple endophenotypes. A facial heatmap is calculated to better visualize the facial deformation, where outward deformation, with respect to the geometric center of the head, were in red color. C-GWAS combines multiple endophenotypes based on single-trait summary statistics of all generated facial endophenotypes. Therefore, given β^{i}_{1}, β^{i}_{2}, …, β^{i}_{n} as betas of n endophenotypes for SNP^{i} in the single-trait GWAS, by decoding and sampling latent datapoints along $\overrightarrow{v^{i}} = \beta^{i}_{1}\overrightarrow{e_1}+\beta^{i}_{2}\overrightarrow{e_2}+ \ldots +\beta^{i}_{n}\overrightarrow{e_n}$, we are able to visualise facial phenotypes associated with SNP^{i} in the C-GWAS results. The range of the sampling is mean ± SD of latent datapoints from the study population. Notably, we only selected nominal-significant ($p<0.05$) endophenotypes for this visualization.

Heritability and genetic correlation analyses

We performed LDSC[7] on the meta-analysis GWASs to estimate the genome-wide heritability for each of the 200 endophenotypes. After the estimation of heritability on facial endophenotypes, we mapped the heritability of the endophenotypes to the human face. Implementation details were described in Supplementary Figure S4 and Figure S5. We applied LDSC to compute the genetic correlation between the 200 endophenotypes. Furthermore, we also estimated genetic correlation (rg) between face shape and brain diseases on a summary level, by using publicly available GWAS summary statistics for 12 diseases and disorders including anxiety disorder[24], attention deficit hyperactivity disorder (ADHD)[25], autism spectrum disorder[26], Alzheimer's disease[27], bipolar disorder[28], cross disorder[29], eating disorder[30], major depressive disorder[31], other Tic Disorders and Tourette syndrome[32], schizophrenia[33], and alcohol use disorders[34]. We applied LDSC[8] to calculate rg for each pair of face and disease GWAS.

Results

We developed an integrated analysis pipeline for exploring the genetic basis of image-based complex traits, which combined i) data-driven AI-based phenotyping to cover image complexity via a large number of endophenotypes, ii) combined genome-wide association studies from a large number of endophenotypes, , iii) AI-based visualization of the identified genetic association effects on the total image, and iv) federated learning to avoid the problem of sharing individual image data in multi-cohort studies, which we applied to the prominent example of the human face (Figure 1).

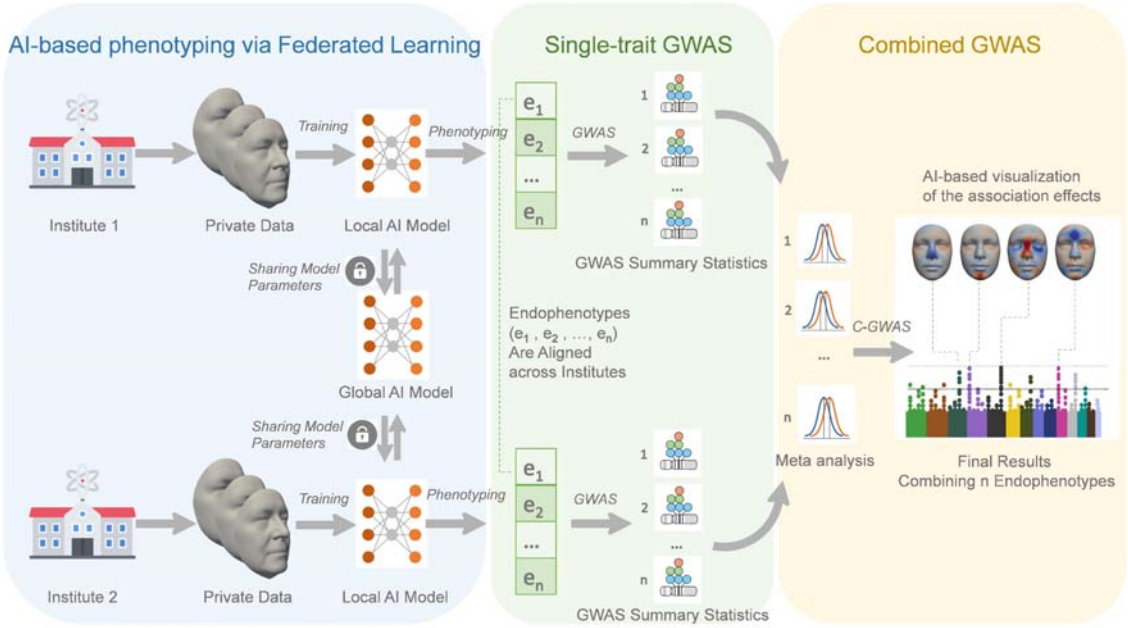

Figure 1: Analysis pipeline to obtain genetic insights into image-based complex traits using the human face as example, which consists of three blocks: 1) AI models were trained via federated learning techniques without inter-institutional sharing individual-level image data and extraction of numerous endophenotypes from the total image.; 2) single-trait GWASs for all AI-derived endophenotypes per each cohort separately; 3) meta-analysis of all single trait GWAS summary statistic data from all endophenotypes and all cohorts and C-GWAS for combing all endophenotypes from all cohorts with AI-based visualization of the revealed association effects on the total image.

Comparison between federating learning and centralized training strategies

AI-based phenotyping requires model-based learning. In traditional CT, images need to be shared between study sites for merging them into one single training session. In practise, due to privacy regulations, inter-institutional image sharing often is not possible, especially for privacy-sensitive images such as facial images. Applying FL to image-based analysis avoids the need of inter-institutional image sharing. To find out if FL performs comparable with CT, we carried out a comparative analysis of image reconstruction with our AI-based facial phenotyping model via FL training and traditional CT (Figure 2).

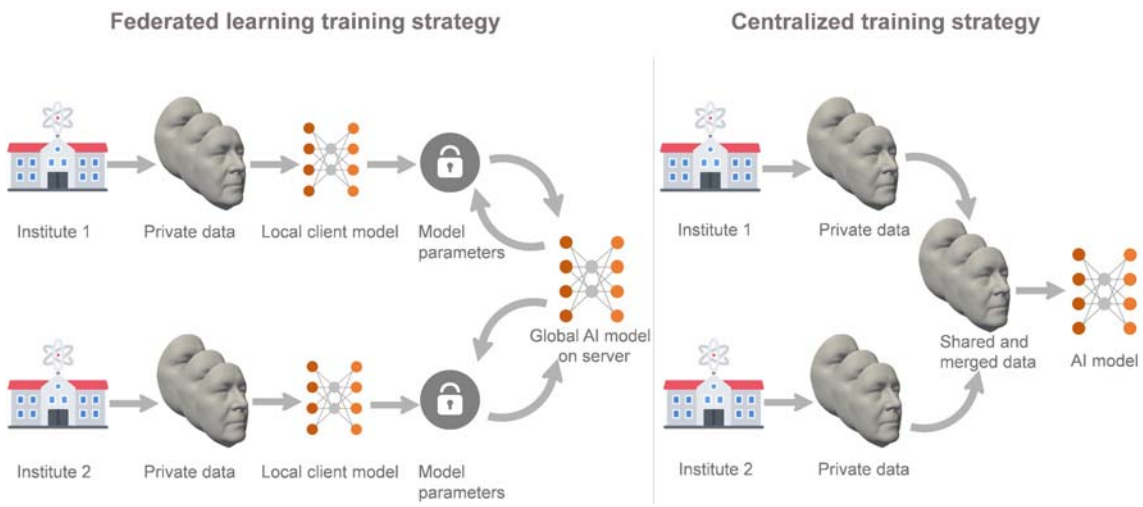

Figure 2: Overview of Federated Learning training strategy and Centralized Training strategy.

We found that our proposed FL training strategy (mean error 0.230 mm ± 0.026 mm) demonstrated similar performance without the need to share images relative to the traditional CT strategy (mean error 0.227 mm ± 0.026 mm) that requires image sharing (Table 1). This indicates that FL is an effective strategy for image-based AI model training, at least when it comes to the face example used here, thereby bypassing the need of sharing individual image data between different institutions. Figure S1 shows details about the reconstruction error in facial heatmaps using an individual from the elderly cohort (RS) and an individual from the children cohort (GenR) as examples. Notably, larger reconstruction errors were found in sharp facial areas such as eyes and nasolabial folds.

Table 1: Mean reconstruction error of FL and CT strategies.

	CT (mm)	**FL (mm)**
RS (N=3,995)	0.232±0.031	0.233±0.031
GenR (N=3,314)	0.220±0.017	0.227±0.018
RS + GenR (N=7,309)	0.227±0.026	0.230±0.026

CT: model was trained by a centralized dataset with merged images from both cohorts RS and GenR,
FL: model was trained via FL settings without creating a centralized dataset across both cohorts.

C-GWAS and replication outcomes

After performing single trait GWASs on all 200 AI-derived facial endophenotypes in both European cohorts separately, and meta-analysis of both cohorts (N=7,309), we carried out C-GWAS to search for genetic associations for all 200 endophenotypes and in both cohorts in a combined way. To demonstrate the phenotypic effects of the obtainable genetic associations, we developed an explainable AI-based facial visualization approach (Figure 3).

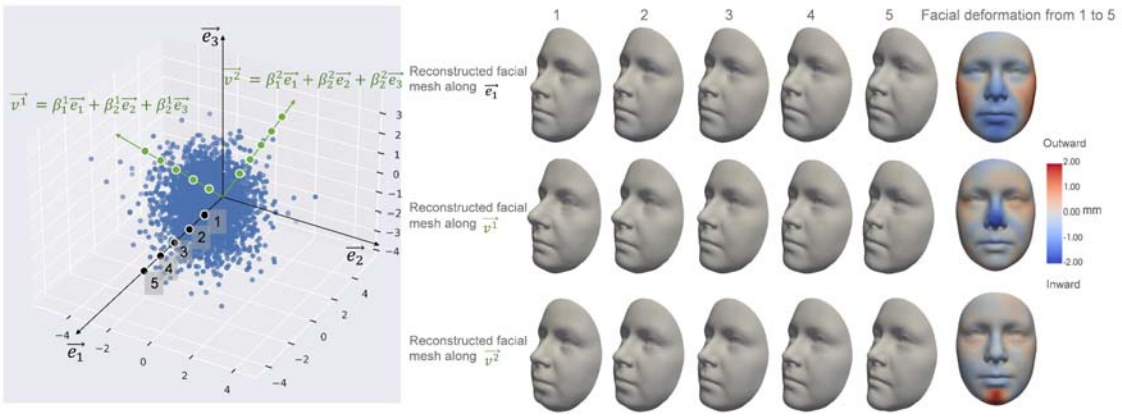

Figure 3: Explainable AI-based facial visualisation of genetic association effects. Endophenotypes were decoded along a vector direction in the endophenotype space (n=3 dimensions for demonstration). Axis $\vec{e_1}$, $\vec{e_2}$, …, $\vec{e_n}$ are basis vectors representing n facial endophenotypes; β^i_1, β^i_2, ..., β^i_n are betas of n endophenotypes for SNP^i in the single-trait GWAS. Vectors $\vec{v^1}$ and $\vec{v^2}$ capture the variance of facial phenotypes associated with SNP^1 and SNP^2 in the C-GWAS results. When sampling along a vector, a series of facial images were reconstructed from the decoder, and outward changes in the reconstructed faces (with respect to the geometric center of the head) were shown as red regions in the heatmap, while blue regions indicate inward changes. Each blue dot in the diagram is a latent data point representing the low-dimensional representation of an input facial image in the endophenotype space.

Based on C-GWAS in both European cohorts combined, we identified 43 genetic loci demonstrating SNPs with genome-wide statistically significant facial associations (Figure 4, Figure S6, Table S1 for more details). Via AI-derived facial heatmaps, we demonstrate the phenotypic effects of the obtained significant associations (Figure 4). An inspection of these 43 loci in all available resources, i.e., a comprehensive review of the previous human face genetic literature (Table S2) and a recently published large-scale face C-GWAS in

Europeans[85], revealed that 12 (28%) of these loci were not previously reported as face loci, marking them as novel genetic face loci we identified here.

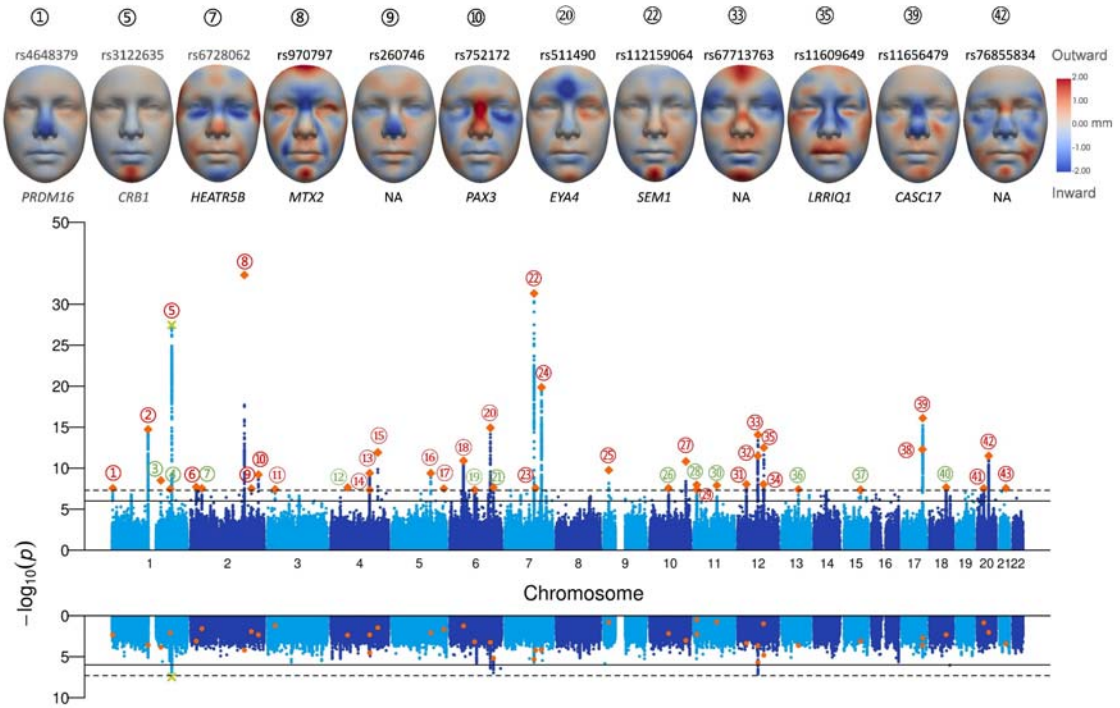

Figure 4: Manhattan plot of C-GWAS outcomes (upper part) from 200 AI-generated facial endophenotypes in two European cohorts combined (N=7,309) and AI-based facial visualization of the association effects of the C-GWAS-derived top SNPs. Green numbers with circles denote novel face loci, while red numbers denote previously known face loci that were confirmed here. Lower part shows results from single-trait GWAS; green-crosses denote SNPs that were significant in both, single-trait GWAS and C-GWAS; red-dots denote SNPs that were only significant in C-GWAS. Red color in the heatmap means outward facial effects w.r.t the geometric center of the head, while blue means inward facial effects.

The AI-derived facial heatmap visualization of the association effects of the 12 novel genetic face loci demonstrated that they are all linked with several facial endophenotypes involving multiple facial regions (Figure 5). This finding underlines the ability of C-GWAS in identifying genetic loci with pleotropic effects.

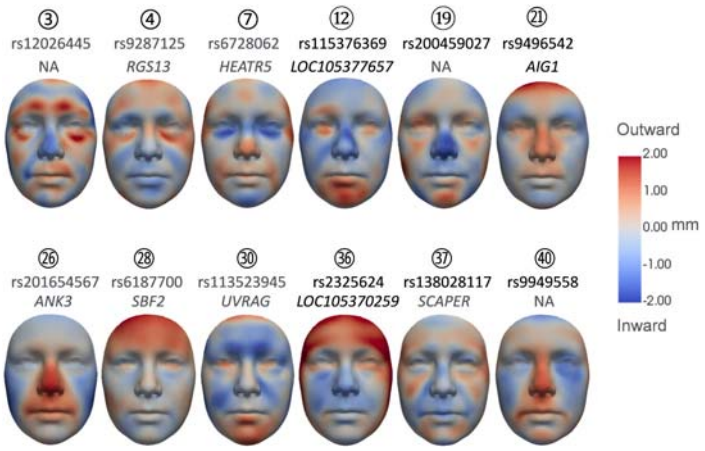

Figure 5: AI-derived facial heatmaps visualizing the association effects of the 12 novel genetic face loci highlighted by the C-GWAS of both European cohorts combined. Red color means outward effects w.r.t the geometric center of the head, while blue means inward effects.

According to annotation information of FUMA[77], the 12 novel face loci are additionally linked to other traits, such as QRS duration (□)[54], brain morphology (□ □)[57], diabetic maculopathy (□)[58], blood cells (□ □)[59], blood pressure measurement (□ □)[60,61], total cholesterol measurement(□)[62], and Alzheimer disease (□)[63].

In the replication analysis using independent European data (N=8,246) [5], 30 (70%) of the 43 loci, of which 2 (7%) were novel loci, demonstrated statistically significant association (i.e., $p < 0.05$), providing confirmatory evidence (Table S1). Figure S3 shows the comparison of facial heatmaps between our findings from the discovery analysis and the replication analysis, providing further evidence that the identified loci exhibit consistent phenotypic patterns in both analyses.

Comparison of C-GWAS outcomes between the two cohorts

Since both discovery cohorts are of largely different age ranges, we additionally performed C-GWAS of the 200 AI-derived facial endophenotypes separately for the children in GenR (N=3,314) and the elderly in RS (N=3,995) (Manhattan plot in Figure S2). With 21 identified

genome-wide significant genetic loci in RS and 18 in GenR, we found about similar number of loci in the elderly cohort as in the children cohort, with only four loci (*CRB1*; *MTX2; EYA4; SEM1*) overlapping between both cohorts.

We further focused on the changes of the statistical power before and after the meta-analysis of both cohorts in C-GWAS. Among the 43 identified loci (Table S1), 37 loci (86%), showed increased power after meta-analysis of both cohorts in C-GWAS (Figure 6). For instance, four loci were genome-wide significant in both cohorts and reached stronger significance in the meta-analysis of both cohorts (Figure 6a), four loci were genome-wide significant in one of the two cohorts and genome-wide significance in the meta-analysis of both cohorts (Figure 6b,c), and five loci were not significant in either of the two cohorts, reached borderline or nominal significance in both cohorts and showed consistent patterns in facial heatmaps across both cohorts, and were genome-wide significant in the meta-analysis of both cohorts (Figure 6d). These findings highlight the benefits of increased sample size in enhancing statistical power through our proposed pipeline despite the different age ranges of the two cohorts.

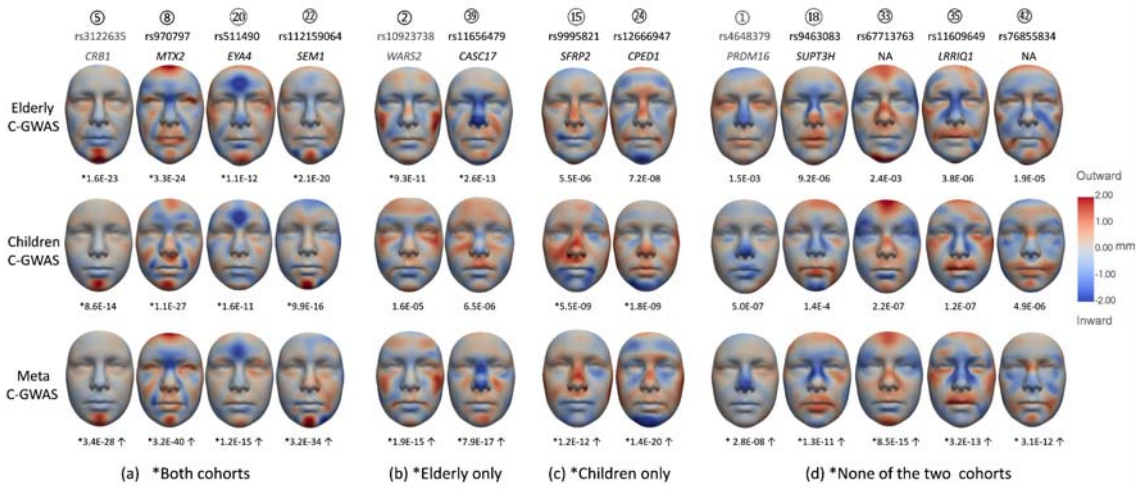

Figure 6: Impact of meta-analysis on loci detection power with C-GWAS. a) Loci genome-wide significant in the two separate C-GWASs in both cohorts; b) Loci genome-wide significant in the

separate C-GWAS of the elderly RS cohort; c) Loci genome-wide significant in the separate C-GWAS of the children GenR cohort; d) Loci genome-wide significant in the meta-analysis C-GWAS of both cohorts but non-significant in the two separate C-GWASs of both cohorts. Red color in the heatmap means outward effects w.r.t the geometric center of the head, while blue means inward effects. Numbers with circles are index of the 43 identified genetic loci in Table S1.

Notably, we also found 5 loci that were genome-wide significant in one of the two cohorts, but not after meta- analysis of both cohorts (Figure 7). The facial visualisation of the associations effects of these 5 loci is largely different between the two cohorts. These findings suggest underlying cohort-specific genetic face effects that may be explained by the large age difference of the participants of both cohorts.

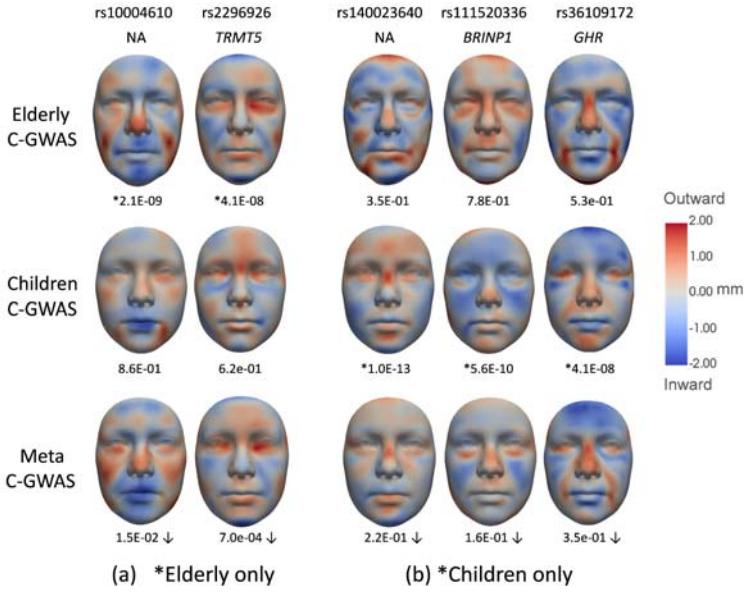

Figure 7: Genetic loci with genome-wide significance identified with separate C-GWASs in one of the two cohorts but not in the C-GWAS of both cohorts combined. a) Genome-wide significant loci in the elderly RS cohort, but not in the children GenR cohort; b) Genome-wide significant loci in the children GenR cohort, but not in the elderly RS cohort. Red color in the heatmap means outward effects w.r.t the geometric center of the head, blue means inward effects.

Heritability analysis outcome

LDSC-based heritability analysis on the single-trait GWAS summary statistics after meta-analysis of the two cohorts was performed and after false discovery rate (FDR) correction for multiple testing, revealed a total of 90 of the 200 (45%) endophenotypes to be statistically significant with the h^2 range from 0.13 to 0.34 (Table S3). This finding is remarkable because only one genetic locus was identified with genome-wide significant association in the single-trait meta-analysis GWAS results (Figure 4). Next, we combined these FDR-significant endophenotypes and mapped them onto the face, revealing combined heritability effects ranging from 0.20 to 0.35. As evident from Figure 8, almost all facial regions are covered. Figures S4 and S5 show implementation details about merging multiple endophenotypes and mapping them back to the face.

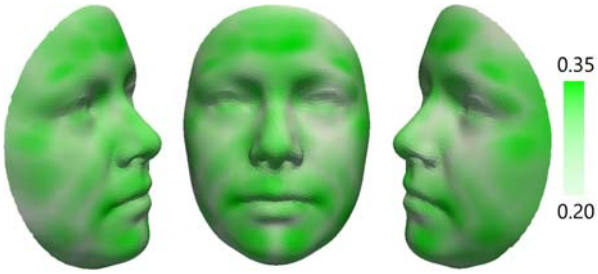

Figure 8: LDSC-based heritability of facial endophenotypes from meta-analysis of single-trait GWAS summary statistics based on the combination of 90 FDR-significant facial endophenotypes.

Genetic correlation outcome

A LDSC-based genetic correlation analysis between the 200 AI-derived facial endophenotypes based on single-trait GWAS summary statistics after meta-analysis of the two cohorts highlighted 21 (10.5%) of the 200 endophenotypes to be genetically correlated with at least one other endophenotype based on FDR correction ($\alpha = 0.05$) performed separately for each endophenotype across its 199 pairwise tests. Among the significant pairs, the absolute genetic correlation values ranged from 0.56 to 0.92 (mean: 0.73). The corresponding phenotypic correlations for these genetically correlated pairs of endophenotypes ranged from 0.05 to 0.67 (mean: 0.23). The existence of high genetic

correlation among facial endophenotypes implies the pleiotropic effects of the respective SNP.

Additionally, we carried out a genetic correlation analysis between the 200 facial endophenotypes and various disease traits (Figure S7). Nominal significance correlations were seen for diseases such as ADHD, bipolar disorder, major depressive disorder, other Tic disorders and Tourette syndrome, schizophrenia, and alcohol use disorders; however, none of them remained statistically significant after multiple testing correction.

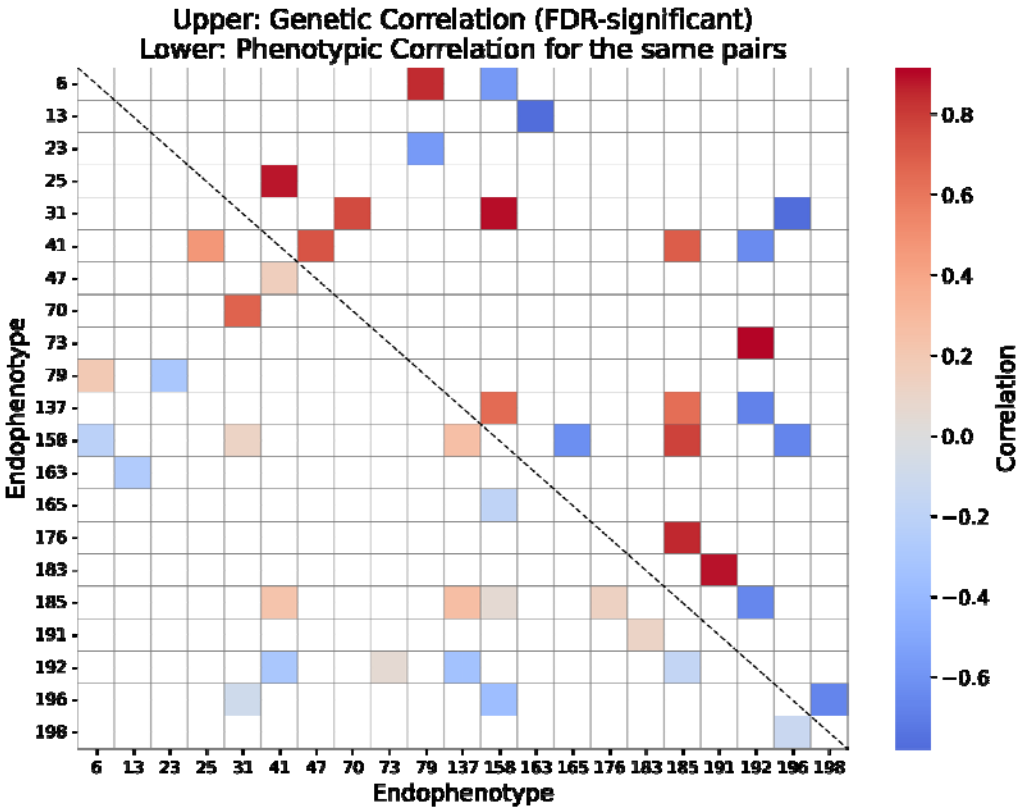

Figure 9: LDSC-based genetic correlation between AI-derived facial endophenotypes with FDR-significant associations, and the corresponding phenotypic correlation.

C-GWAS comparison of AI-derived phenotypes and landmark-based phenotypes

To explore the impact of our novel AI-based versus previously used landmark-based facial phenotyping strategies on C-GWAS outcomes, we performed on the same 3,761 RS samples

a C-GWAS comparison between AI-derived and landmark-based facial phenotypes we had used previously[9]. The QQ plot indicates that the C-GWAS based on AI-derived facial phenotypes exhibits higher statistical power than that of landmark-based facial phenotypes (Figure 10). In particular, the C-GWAS based on AI-derived facial phenotypes identified 13 significant loci, whereas that from landmark-based phenotypes revealed 10 significant loci, with only three loci overlapping between the two (Figure 10). This limited overlap in the same samples suggests that the two phenotyping approaches capture largely distinct aspects of facial shape variation with different underlying genetic basis.

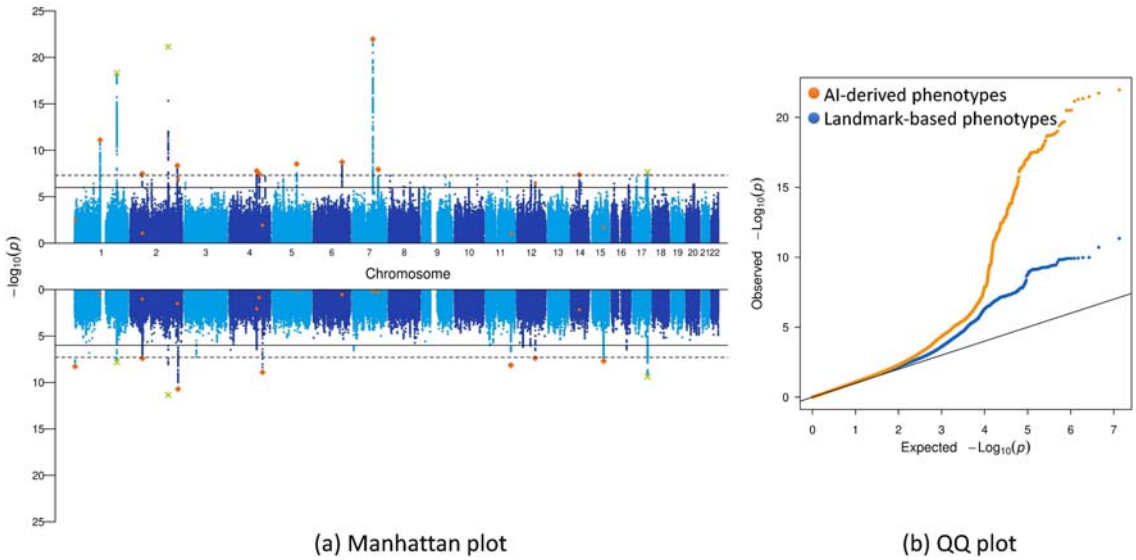

Figure 10: C-GWAS comparison of AI-derived and landmark-based facial phenotypes. a) Manhattan plot: AI-derived (top) vs. landmark-based (bottom) C-GWAS outcomes. Green-crosses denote SNPs that were significant in both analyses. Red-dots denote SNPs that were only significant in one of the two analyses; b) QQ plot: AI-derived (orange) vs. landmark-based (blue) results.

Discussion

In this study, we developed an innovative, comprehensive, integrated AI-driven analysis pipeline to obtain genetic insights into image-based complex traits from multi-cohort data. The proposed pipeline is fully data-driven; hence, no a-priory knowledge is required.

Compared to landmark-based phenotypes, which rely on prior knowledge, our data-driven approach demonstrated higher statistical power and identified more novel genetic loci. Through training with the FL strategy, we achieve endophenotype alignment across different cohorts with similar accuracy compared to the traditional CT approach. Because FL in contrast to CT works without sharing individual image data, applying FL represents a strong advantage given that image data can be sensitive and often cannot be shared across institutions. Inter-institutional multi-cohort studies are required to increase sample size when studying complex trait characterized by small effect size. Thus, by applying our approach, inter-institutional multi-cohort studies can be carried out without sharing individual image data. Our final genetic analysis combines the GWASs of the 200 AI-derived endophenotype using the C-GWAS approach to enhance statistical power. Our AI-driven visualization of the genetic associations allows interpreting them via illustrating their effects on the human face.

The effectiveness of our analysis pipeline was demonstrated with the example of 3D facial images. Applying the pipeline to a moderate sample size (N=7,312) from two cohorts, we identified 43 genetic loci, of which 31 represent previously identified genetic face loci. Notably, for known genomic loci, the highlighted patterns in the AI-based facial heatmaps align remarkably well with the effects described for these loci in the literature[5]. For example, *PRDM16*, *PAX3,* and *CRB1* were previously described to be associated with the nose apex[5,37], nasal bridge[9] and chin dimples[5], respectively, which agrees with our AI-based facial visualization results of the genetic associations we re-discovered for these loci in our data. The 31 re-discovered identified loci also showed associations with various health conditions, such as lung function (*ALX1, CASC17*)[40,41], asthma (*CRB1*)[42], coronary artery disease (*CRB1*)[43], liver enzyme levels (*SEM1*)[55], insomnia (*CRB1*)[44,45], Alzheimer's disease (*EYA4*)[51], brain morphology (*PAX3, SEM1, ALX1*)[52,53]. Next to providing confirmatory evidence for these previously identified loci, our results demonstrates that our approach using AI-based

facial phenotypes is able to re-identify loci previously discovered by use of substantially different ways of facial phenotyping used previously.

In addition to re-identifying 31 previously known genetic face loci, our approach also highlighted 12 genetic loci that were unknown yet to be involved in facial variation. This finding is remarkable given the moderate sample size we used, which was smaller than in some previous studies. Our high re-discovery rate of 72% provides general confidence that the newly identified loci reflect true positive findings. Moreover, empirical confirmation for 30 of them was provided by our replication analysis in independent European samples. The 12 novel loci are linked to other traits, such as QRS duration (*HEATR5B*)[54], brain morphology (*HEATR5B, ANK3*)[57], total cholesterol measurement (*UVRAG*)[62], and Alzheimer disease (*SCAPER*)[63]. These traits are known to be associated with the human face[64,66,67,68], and therefore, these findings further justify the robustness of our AI-based approach and validate the identification of these novel loci. Functional studies and experimental validation are needed to explore how these genetic variants influence both facial traits and their associated phenotypic pathways.

C-GWAS outperformed single-trait GWAS

All previous image-based GWAS studies used the traditional single-trait GWAS approach on endophenotypes derived from autoencoders, e.g. single-endophenotype GWAS on brain[14] and cardiac endophenotypes[15]. Specifically, one study demonstrated that facial mesh endophenotypes generated by autoencoders have limited performance in single-endophenotype GWAS[80]. Our results not only confirm this limitation but also provide insight into how it can be addressed using C-GWAS, which leverages multiple endophenotypes. As shown in Figure 4, in our dataset, single-endophenotype GWAS identified significantly fewer hits, i.e., only one, compared to the many multi-endophenotype GWAS hits (43) achieved

with C-GWAS. The underlying reason for this phenomenon partly lies in the unsupervised nature of the applied AI models, which cannot ensure that the derived endophenotypes (i.e., basis vectors $\overrightarrow{e_1}, \overrightarrow{e_2}$, etc. in Figure 3) correspond to facial phenotypes that are strongly associated with a SNP. For example, the nose apex is known to be associated with *PRDM16*[5,37], but we failed to identify it among any of the 200 endophenotypes in our single-endophenotype GWAS. This is because the variability of the nose apex was precisely captured by another vector instead of any of the 200 basis vectors. From this perspective, it is important to search—guided by C-GWAS—for additional facial phenotypes represented by other vectors beyond the 200 basis vectors. Moreover, given that previous face GWASs and C-GWASs based on other ways of facial phenotyping have already discovered hundreds of genetic loci [5,6,9,37,38,81,82,83,84], of which 31 were re-discovered here, our approach of AI-based phenotyping coupled with C-GWASs allows to unveil new genetic loci involved in facial shapes because they were likely not covered by the previously used facial phenotypes. Thus, for the example of the human face at least, our novel approach reflects a synergistic approach to understand the genetic architecture underlying variation of facial shape more completely.

FL for alignment of endophenotypes across cohorts

The alignment of endophenotypes across multiple cohorts poses a significant challenge in data-driven AI-based phenotyping due to variations in data distributions. Traditionally, this challenge has been addressed through centralized training strategies, where data from multiple cohorts are shared and combined into a single source for model training. However, this approach is not in line with the practise of GWAS meta-analysis in multi-cohort studies needed to achieve larger sample size, where researchers typically prefer to share GWAS summary statistics data between collaborating institutions instead of individual data.

Moreover, sharing individual-level images, especially those that allow individualization such as facial images, often have strong hurdles in sharing across institutions and countries, which sometimes cannot be solved and thus such image data cannot be shared. To overcome this limitation, we integrated an alternative training strategy: federated learning (FL) in our analysis pipeline. Our proposed FL strategy has demonstrated comparable performance to traditional CT in deriving aligned endophenotypes across different cohorts, without the need for image sharing between different sites. This property of our pipeline facilitates collaborative C-GWAS efforts involving multi-cohort data from different institutions.

Comparison between children and elderly cohorts

The efficacy of our FL strategy was demonstrated through a multi-cohort study of an elderly cohort (mean age=69.6 years, SD=9.5 years) and a children cohort (mean age=10.4 years, SD=1.5 years). Despite the significant age difference between the participants of these two cohorts, the combination of data in the joined C-GWAS led to an improvement in power for most of the identified genetic loci. For instance, *PRDM16*, *CRB1*, *MTX2*, and *EYA4* exhibited similar patterns in facial heatmaps across the two cohorts, and their statistical power was enhanced after the meta-analysis (Figure 6). This highlights the success of our proposed pipeline in conducting GWAS meta-analysis across multiple cohorts.

However, we also observed a decrease in power for five genetic loci after the meta-analysis. For example, rs2296926 (*TRMT5*, p-value: 4.1E-08 in RS), rs111520336 (*BRINP1*, p-value: 5.6e-10 in GenR) and rs36109172 (*GHR*, p-value: 4.1E-08 in GenR) showed different facial patterns between the two cohorts (Figure 7). While they were genome-wide significant in either the elderly cohort or the children cohort alone with separate C-GWASs, they did not retain statistical significance after meta-analysis in the combined C-GWAS. We hypothesize that this discrepancy may be attributed to the large difference in age between the

two cohorts. This hypothesis is substantiated by some of the identified loci. For instance, *BRINP1* is strongly associated with body height, adolescent idiopathic scoliosis, and cognitive function[69,70,71], while *TRMT5* is highly related to bone density[72], and *GHR* is closely linked to body height and growth hormone receptor measurement[69,74]. These traits, like facial morphology, are age-dependent and likely undergo dynamic changes across the lifespan. Age-related differences may alter the expression or relevance of genetic loci, potentially shifting their impact on facial phenotypes at different stages of life.

Strengths and limitations

Different from other existing multi-trait GWAS pipelines[5], our pipeline that finally uses C-GWAS, retains conventional single-trait GWAS summary statistics, thus offering the opportunity to perform genetic correlation analysis with other phenotypes. Our AI-based facial endophenotypes derived from population cohorts demonstrate genetic correlations with various diseases; however, the obtained correlation is weak and lacks statistical significance. This result is in line with prior findings about the weak genetic correlation between diseases and normal facial traits[64]. Moreover, we identified high genetic correlation among the 200 AI-derived facial endophenotypes, which indicates the pleiotropic effects of a SNP on multiple endophenotypes. Therefore, applying C-GWAS to combine the effects of multiple endophenotypes is crucial to improve the power, especially when sample size is modest. As shown in Figure 4, multi-endophenotype GWAS (43 hits) facilitated through the C-GWAS approach, yielded significantly more hits than the single-endophenotype GWAS (one hit).

A notable limitation of this study lies in the significant disparity in age distribution between the two cohorts, which likely restricted the extent of power enhancement through meta-analysis. Nevertheless, we anticipate that the performance of our pipeline would be notably improved in scenarios where the age distribution across cohorts exhibits greater similarity. Based on the compelling evidence from this study, we anticipate that our novel

analysis pipeline will emerge as a crucial tool for facilitating inter-institutional collaborations involving multiple cohorts to unveil the genetic basis of image-based complex phenotypes, while effectively addressing data privacy concerns by preventing inter-institutional sharing of individual images.

Acknowledgement

We thank all cohort participants or their legal representatives for their participation.

Data availability

The links to data and scripts will be released after acceptance.

Reference:

[1] A. Hofman, S. D. Murad, C. M. van Duijn, O. H. Franco, A. Goedegebure, M. Arfan Ikram, C. C. Klaver, T. E. Nijsten, R. P. Peeters, B. H. C. Stricker, et al., The Rotterdam Study: 2014 objectives and design update. *European journal of epidemiology*. 2013;28():889-926.
[2] M. Kayser, F. Liu, A. C. J. Janssens, F. Rivadeneira, O. Lao, K. van Duijn, M. Vermeulen, P. Arp, M. M. Jhamai, W. F. van IJcken, et al., Three genome-wide association studies and a linkage analysis identify HERC2 as a human iris color gene. *The American Journal of Human Genetics*. 2008;82(2):411-423.
[3] 1. G. P. Consortium, et al., An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012;491(7422):56.
[4] S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M. A. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. De Bakker, M. J. Daly, et al., PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American journal of human genetics*. 2007;81(3):559-575.
[5] J. D. White, K. Indencleef, S. Naqvi, R. J. Eller, H. Hoskens, J. Roosenboom, M. K. Lee, J. Li, J. Mohammed, S. Richmond, et al., Insights into the genetic architecture of the human face. *Nature genetics*. 2021;53(1):45-53.
[6] Z. Xiong, X. Gao, Y. Chen, Z. Feng, S. Pan, H. Lu, A. G. Uitterlinden, T. Nijsten, A. Ikram, F. Rivadeneira, et al., Combining genome-wide association studies highlight novel loci involved in human facial variation. *Nature Communications*. 2022;13(1):7832.
[7] B. Bulik-Sullivan, H. K. Finucane, V. Anttila, A. Gusev, F. R. Day, P. Loh, R. Consortium, P. G. Consortium, G. C. f. A. N. o. t. W. T. C. C. C. 3, L. Duncan, et al., An atlas of genetic correlations across human diseases and traits. *Nature genetics*. 2015;47(11):1236-1241.

[8] B. K. Bulik-Sullivan, P. Loh, H. K. Finucane, S. Ripke, J. Yang, S. W. G. o. t. P. G. Consortium, N. Patterson, M. J. Daly, A. L. Price, B. M. Neale. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature genetics*. 2015;47(3):291-295.
[9] Z. Xiong, G. Dankova, L. J. Howe, M. K. Lee, P. G. Hysi, M. A. De Jong, G. Zhu, K. Adhikari, D. Li, Y. Li, et al., Novel genetic loci affecting facial shape variation in humans. *elife*. 2019;8():e49898.
[10] P. R. Jansen, M. Nagel, K. Watanabe, Y. Wei, J. E. Savage, C. A. de Leeuw, M. P. van den Heuvel, S. van der Sluis, D. Posthuma. Genome-wide meta-analysis of brain volume identifies genomic loci and genes shared with intelligence. *Nature communications*. 2020;11(1):5606.
[11] J. L. Stein, X. Hua, S. Lee, A. J. Ho, A. D. Leow, A. W. Toga, A. J. Saykin, L. Shen, T. Foroud, N. Pankratz, et al., Voxelwise genome-wide association study (vGWAS). *neuroimage*. 2010;53(3):1160-1174.
[12] B. Zhao, T. Luo, T. Li, Y. Li, J. Zhang, Y. Shan, X. Wang, L. Yang, F. Zhou, Z. Zhu, et al., Genome-wide association analysis of 19,629 individuals identifies variants influencing regional brain volumes and refines their genetic co-architecture with cognitive and mental health traits. *Nature genetics*. 2019;51(11):1637-1644.
[13] P. Claes, J. Roosenboom, J. D. White, T. Swigut, D. Sero, J. Li, M. K. Lee, A. Zaidi, B. C. Mattern, C. Liebowitz, et al., Genome-wide mapping of global-to-local genetic effects on human facial shape. *Nature genetics*. 2018;50(3):414-423.
[14] K. Patel, Z. Xie, H. Yuan, S. M. S. Islam, Y. Xie, W. He, W. Zhang, A. Gottlieb, H. Chen, L. Giancardo, et al., Unsupervised deep representation learning enables phenotype discovery for genetic association studies of brain imaging. *Communications Biology*. 2024;7(1):414.
[15] R. Bonazzola, E. Ferrante, N. Ravikumar, Y. Xia, B. Keavney, S. Plein, T. Syeda-Mahmood, A. F. Frangi. Unsupervised ensemble-based phenotyping enhances discoverability of genes related to left-ventricular morphology. *Nature Machine Intelligence*. 2024;6(3):291-306.
[16] Z. Xie, T. Zhang, S. Kim, J. Lu, W. Zhang, C. Lin, M. Wu, A. Davis, R. Channa, L. Giancardo, et al., iGWAS: Image-based genome-wide association of self-supervised deep phenotyping of retina fundus images. *PLoS genetics*. 2024;20(5):e1011273.
[17] M. A. Ikram, G. Brusselle, M. Ghanbari, A. Goedegebure, M. K. Ikram, M. Kavousi, B. C. Kieboom, C. C. Klaver, R. J. de Knegt, A. I. Luik, et al., Objectives, design and main findings until 2020 from the Rotterdam Study. *European journal of epidemiology*. 2020;35():483-517.
[18] V. W. Jaddoe, J. P. Mackenbach, H. A. Moll, E. A. Steegers, H. Tiemeier, F. C. Verhulst, J. C. Witteman, A. Hofman. The Generation R Study: design and cohort profile. *European journal of epidemiology*. 2006;21():475-484.
[19] V. W. Jaddoe, C. M. van Duijn, A. J. van der Heijden, J. P. Mackenbach, H. A. Moll, E. A. Steegers, H. Tiemeier, A. G. Uitterlinden, F. C. Verhulst, A. Hofman. The Generation R Study: design and cohort update 2010. *European journal of epidemiology*. 2010;25():823-841.
[20] X. Liu, M. Kayser, S. Kushner, H. Tiemeier, F. Rivadeneira, V. Jaddoe, W. Niessen, E. Wolvius, G. Roshchupkin. Association between prenatal alcohol exposure and children's facial shape: a prospective population-based cohort study. *Human Reproduction*. 2023;38(5):961-972.
[21] S. Gong, L. Chen, M. Bronstein, S. Zafeiriou. SpiralNet++: A Fast and Highly Efficient Mesh Convolution Operator. In: *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*. 2019:4141–4148. doi:10.1109/ICCVW.2019.00509.
[22] Wikipedia contributors. *Federated learning — Wikipedia, The Free Encyclopedia*. Wikipedia, The Free Encyclopedia. 2024. Available from: https://en.wikipedia.org/wiki/Federated_learning. [Accessed 2 January 2024].
[23] G. V. Roshchupkin, H. Adams, M. Vernooij, A. Hofman, C. Van Duijn, M. Ikram, W. Niessen. HASE: Framework for efficient high-dimensional association analyses. *Scientific reports*. 2016;6(1):36076.

[24] A. J. Forstner, S. Awasthi, C. Wolf, E. Maron, A. Erhardt, D. Czamara, E. Eriksson, C. Lavebratt, C. Allgulander, N. Friedrich, et al., Genome-wide association study of panic disorder reveals genetic overlap with neuroticism and depression. *Molecular psychiatry*. 2021;26(8):4179-4190.
[25] J. Martin, R. K. Walters, D. Demontis, M. Mattheisen, S. H. Lee, E. Robinson, I. Brikell, L. Ghirardi, H. Larsson, P. Lichtenstein, et al., A genetic investigation of sex bias in the prevalence of attention-deficit/hyperactivity disorder. *Biological psychiatry*. 2018;83(12):1044-1053.
[26] J. Grove, S. Ripke, T. D. Als, M. Mattheisen, R. K. Walters, H. Won, J. Pallesen, E. Agerbo, O. A. Andreassen, R. Anney, et al., Identification of common genetic risk variants for autism spectrum disorder. *Nature genetics*. 2019;51(3):431-444.
[27] I. E. Jansen, J. E. Savage, K. Watanabe, J. Bryois, D. M. Williams, S. Steinberg, J. Sealock, I. K. Karlsson, S. Hägg, L. Athanasiu, et al., Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nature genetics*. 2019;51(3):404-413.
[28] N. Mullins, A. J. Forstner, K. S. O'Connell, B. Coombes, J. R. Coleman, Z. Qiao, T. D. Als, T. B. Bigdeli, S. Børte, J. Bryois, et al., Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nature genetics*. 2021;53(6):817-829.
[29] P. H. Lee, V. Anttila, H. Won, Y. A. Feng, J. Rosenthal, Z. Zhu, E. M. Tucker-Drob, M. G. Nivard, A. D. Grotzinger, D. Posthuma, et al., Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. *Cell*. 2019;179(7):1469-1482.
[30] H. J. Watson, Z. Yilmaz, L. M. Thornton, C. Hübel, J. R. Coleman, H. A. Gaspar, J. Bryois, A. Hinney, V. M. Leppä, M. Mattheisen, et al., Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. *Nature genetics*. 2019;51(8):1207-1214.
[31] N. R. Wray, S. Ripke, M. Mattheisen, M. Trzaskowski, E. M. Byrne, A. Abdellaoui, M. J. Adams, E. Agerbo, T. M. Air, T. M. Andlauer, et al., Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature genetics*. 2018;50(5):668-681.
[32] D. Yu, J. H. Sul, F. Tsetsos, M. S. Nawaz, A. Y. Huang, I. Zelaya, C. Illmann, L. Osiecki, S. M. Darrow, M. E. Hirschtritt, et al., Interrogating the genetic determinants of Tourette's syndrome and other tic disorders through genome-wide association studies. *American Journal of Psychiatry*. 2019;176(3):217-227.
[33] S. Ripke, C. O'dushlaine, K. Chambert, J. L. Moran, A. K. Kähler, S. Akterin, S. E. Bergen, A. L. Collins, J. J. Crowley, M. Fromer, et al., Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nature genetics*. 2013;45(10):1150-1159.
[34] S. Sanchez-Roige, A. A. Palmer, P. Fontanillas, S. L. Elson, t. S. U. D. W. G. o. t. P. G. C. 23andMe Research Team, M. J. Adams, D. M. Howard, H. J. Edenberg, G. Davies, R. C. Crist, et al., Genome-wide association study meta-analysis of the alcohol use disorders identification test (AUDIT) in two population-based cohorts. *American Journal of Psychiatry*. 2019;176(2):107-118.
[35] M. N. Kooijman, C. J. Kruithof, C. M. van Duijn, L. Duijts, O. H. Franco, M. H. van IJzendoorn, J. C. de Jongste, C. C. Klaver, A. van der Lugt, J. P. Mackenbach, et al., The Generation R Study: design and cohort update 2017. *European journal of epidemiology*. 2016;31():1243-1264.
[36] E. W. Troe, H. Raat, V. W. Jaddoe, A. Hofman, E. A. Steegers, F. C. Verhulst, J. C. Witteman, J. P. Mackenbach. Smoking during pregnancy in ethnic populations: the Generation R study. *Nicotine & tobacco research*. 2008;10(8):1373-1384.
[37] H. Hoskens, D. Liu, S. Naqvi, M. K. Lee, R. J. Eller, K. Indencleef, J. D. White, J. Li, M. H. Larmuseau, G. Hens, et al., 3D facial phenotyping by biometric sibling matching used in contemporary genomic methodologies. *PLoS genetics*. 2021;17(5):e1009528.
[38] B. Bonfante, P. Faux, N. Navarro, J. Mendoza-Revilla, M. Dubied, C. Montillot, E. Wentworth, L. Poloni, C. Varón-González, P. Jones, et al., A GWAS in Latin Americans identifies novel face

shape loci, implicating VPS13B and a Denisovan introgressed region in facial variation. *Science advances*. 2021;7(6):eabc6160.
[39] X. Liu, B. Li, E. E. Bron, W. J. Niessen, E. B. Wolvius, G.V. Roshchupkin. Projection-wise disentangling for fair and interpretable representation learning: Application to 3d facial shape analysis. In: *24th International Conference of Medical Image Computing and Computer Assisted Intervention (MICCAI)*. 2021;Part V 24:814-823. doi.org/10.1007/978-3-030-87240-3_78.

[40] G. Kichaev, G. Bhatia, P. Loh, S. Gazal, K. Burch, M. K. Freund, A. Schoech, B. Pasaniuc, A. L. Price. Leveraging polygenic functional enrichment to improve GWAS power. *The American Journal of Human Genetics*. 2019;104(1):65-75.
[41] D. B. Hancock, M. S. Artigas, S. A. Gharib, A. Henry, A. Manichaikul, A. Ramasamy, D. W. Loth, M. Imboden, B. Koch, W. L. McArdle, et al., Genome-wide joint meta-analysis of SNP and SNP-by-smoking interaction identifies novel loci for pulmonary function. *PLoS genetics*. 2012;8(12):e1003098.
[42] A. L. Marat, P. S. McPherson. Variants of DENND1B associated with asthma in children. *The New England journal of medicine*. 2010;363(10):988-989.
[43] C. Tcheandjieu, X. Zhu, A. T. Hilliard, S. L. Clarke, V. Napolioni, S. Ma, K. M. Lee, H. Fang, F. Chen, Y. Lu, et al., Large-scale genome-wide association study of coronary artery disease in genetically diverse populations. *Nature medicine*. 2022;28(8):1679-1692.
[44] K. Watanabe, P. R. Jansen, J. E. Savage, P. Nandakumar, X. Wang, D. A. Hinds, J. Gelernter, D. F. Levey, R. Polimanti, et al., Genome-wide meta-analysis of insomnia prioritizes genes associated with metabolic and psychiatric pathways. *Nature genetics*. 2022;54(8):1125-1132.
[45] E. M. Byrne, P. R. Gehrman, S. E. Medland, D. R. Nyholt, A. C. Heath, P. A. Madden, I. B. Hickie, C. M. Van Duijn, A. K. Henders, G. W. Montgomery, et al., A genome-wide association study of sleep habits and insomnia. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*. 2013;162(5):439-451.
[46] M. Ikeda, A. Takahashi, Y. Kamatani, Y. Momozawa, T. Saito, K. Kondo, A. Shimasaki, K. Kawase, T. Sakusabe, Y. Iwayama, et al., Genome-wide association study detected novel susceptibility genes for schizophrenia and shared trans-populations/diseases genetic effect. *Schizophrenia bulletin*. 2019;45(4):824-834.
[47] A. F. Pardiñas, P. Holmans, A. J. Pocklington, V. Escott-Price, S. Ripke, N. Carrera, S. E. Legge, S. Bishop, D. Cameron, M. L. Hamshere, et al., Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nature genetics*. 2018;50(3):381-389.
[48] Y. R. Li, J. Li, S. D. Zhao, J. P. Bradfield, F. D. Mentch, S. M. Maggadottir, C. Hou, D. J. Abrams, D. Chang, F. Gao, et al., Meta-analysis of shared genetic architecture across ten pediatric autoimmune diseases. *Nature medicine*. 2015;21(9):1018-1027.
[49] G. Fatemifar, C. J. Hoggart, L. Paternoster, J. P. Kemp, I. Prokopenko, M. Horikoshi, V. J. Wright, J. H. Tobias, S. Richmond, A. I. Zhurov, et al., Genome-wide association study of primary tooth eruption identifies pleiotropic loci associated with height and craniofacial distances. *Human Molecular Genetics*. 2013;22(18):3807-3817.
[50] S. P. Hagenaars, W. D. Hill, S. E. Harris, S. J. Ritchie, G. Davies, D. C. Liewald, C. R. Gale, D. J. Porteous, I. J. Deary, R. E. Marioni. Genetic prediction of male pattern baldness. *PLoS genetics*. 2017;13(2):e1006594.
[51] E. O. Adewuyi, E. K. O'Brien, D. R. Nyholt, T. Porter, S. M. Laws. A large-scale genome-wide cross-trait analysis reveals shared genetic architecture between Alzheimer's disease and gastrointestinal tract disorders. *Communications Biology*. 2022;5(1):691.

[52] A. A. Shadrin, T. Kaufmann, D. van der Meer, C. E. Palmer, C. Makowski, R. Loughnan, T. L. Jernigan, T. M. Seibert, D. J. Hagler, O. B. Smeland, et al., Vertex-wise multivariate genome-wide association study identifies 780 unique genetic loci associated with cortical morphology. *Neuroimage*. 2021;244():118603.
[53] E. Hofer, G. V. Roshchupkin, H. H. Adams, M. J. Knol, H. Lin, S. Li, H. Zare, S. Ahmad, N. J. Armstrong, C. L. Satizabal, et al., Genetic correlations and genome-wide associations of cortical structure in general population samples of 22,824 adults. *Nature communications*. 2020;11(1):4796.
[54] P. Van der Harst, J. van Setten, N. Verweij, G. Vogler, L. Franke, M. T. Maurano, X. Wang, I. Mateo Leach, M. Eijgelsheim, N. Sotoodehnia, et al., 52 genetic loci influencing myocardial mass. *Journal of the American College of Cardiology*. 2016;68(13):1435-1448.
[55] R. Pazoki, M. Vujkovic, J. Elliott, E. Evangelou, D. Gill, M. Ghanbari, P. J. van der Most, R. C. Pinto, M. Wielscher, M. Farlik, et al., Genetic analysis in European ancestry individuals identifies 517 loci associated with liver enzymes. *Nature communications*. 2021;12(1):2579.
[56] M. Vujkovic, J. M. Keaton, J. A. Lynch, D. R. Miller, J. Zhou, C. Tcheandjieu, J. E. Huffman, T. L. Assimes, K. Lorenz, X. Zhu, et al., Discovery of 318 new risk loci for type 2 diabetes and related vascular outcomes among 1.4 million participants in a multi-ancestry meta-analysis. *Nature genetics*. 2020;52(7):680-691.
[57] D. Van Der Meer, T. Kaufmann, A. A. Shadrin, C. Makowski, O. Frei, D. Roelfs, J. Monereo-Sánchez, D. E. Linden, J. Rokicki, D. Alnæs, et al., The genetic architecture of human cortical folding. *Science advances*. 2021;7(51):eabj9446.
[58] W. Meng, B. W. Chan, C. Ezeonwumelu, H. L. Hébert, A. Campbell, V. Soler, C. N. Palmer. A genome-wide association study implicates that the TTC39C gene is associated with diabetic maculopathy with decreased visual acuity. *Ophthalmic Genetics*. 2019;40(3):252-258.
[59] M. Chen, L. M. Raffield, A. Mousas, S. Sakaue, J. E. Huffman, A. Moscati, B. Trivedi, T. Jiang, P. Akbari, D. Vuckovic, et al., Trans-ethnic and ancestry-specific blood-cell genetics in 746,667 individuals from 5 global populations. *Cell*. 2020;182(5):1198-1213.
[60] A. Giri, J. N. Hellwege, J. M. Keaton, J. Park, C. Qiu, H. R. Warren, E. S. Torstenson, C. P. Kovesdy, Y. V. Sun, O. D. Wilson, et al., Trans-ethnic association study of blood pressure determinants in over 750,000 individuals. *Nature gen61etics*. 2019;51(1):51-62.
[61] E. Evangelou, H. R. Warren, D. Mosen-Ansorena, B. Mifsud, R. Pazoki, H. Gao, G. Ntritsos, N. Dimou, C. P. Cabrera, I. Karaman, et al., Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nature genetics*. 2018;50(10):1412-1425.
[62] S. Lee, J. Choi, B. Park, M. Cha, K. Hong, D. Jung. Dyslipidaemia—Genotype Interactions with Nutrient Intake and Cerebro-Cardiovascular Disease. *Biomedicines*. 2022;10(7):1615.
[63] C. Gouveia, E. Gibbons, N. Dehghani, J. Eapen, R. Guerreiro, J. Bras. Genome-wide association of polygenic risk extremes for Alzheimer's disease in the UK Biobank. *Scientific Reports*. 2022;12(1):8404.
[64] S. Naqvi, Y. Sleyp, H. Hoskens, K. Indencleef, J. P. Spence, R. Bruffaerts, A. Radwan, R. J. Eller, S. Richmond, M. D. Shriver, et al., Shared heritability of human face and brain shape. *Nature Genetics*. 2021;53(6):830-839.
[65] H. R. Roth, Y. Cheng, Y. Wen, I. Yang, Z. Xu, Y. Hsieh, K. Kersten, A. Harouni, C. Zhao, K. Lu, et al., Nvidia flare: Federated learning from simulation to real-world. *arXiv preprint arXiv:2210.13291*. 2022;():.
[66] M. Milanic, A. Bjorgan, M. Larsson, P. Marraccini, T. Strömberg, L. Lyngsnes Randeberg. Hyperspectral imaging for detection of cholesterol in human skin. In: *Optical Diagnostics and Sensing XV: Toward Point-of-Care Diagnostics*. 2015;9332:130-141. SPIE.
[67] E. Conte, M. Chiesa, E. Melotti, D. Marchetti, M. Schillaci, S. Mushtaq, A. Baggiano, G. Pontone, A. Bartorelli, D. Andreini. Artificial intelligence analysis applied to facial features for

prediction of high-risk coronary atherosclerosis at cardiac CT: the LOMBROSO Trial. *European Heart Journal*. 2023;44(Supplement_2):ehad655.2460.
[68] C. Chien, J. Sung, C. Wang, C. Yen, Y. Yang. Analyzing Facial Asymmetry in Alzheimer's Dementia Using Image-Based Technology. *Biomedicines*. 2023;11(10):2802.
[69] L. Yengo, S. Vedantam, E. Marouli, J. Sidorenko, E. Bartell, S. Sakaue, M. Graff, A. U. Eliasen, Y. Jiang, S. Raghavan, et al., A saturated map of common genetic variants associated with human height. *Nature*. 2022;610(7933):704-712.
[70] J. Liu, Y. Zhou, S. Liu, X. Song, X. Yang, Y. Fan, W. Chen, Z. C. Akdemir, Z. Yan, Y. Zuo, et al., The coexistence of copy number variations (CNVs) and single nucleotide polymorphisms (SNPs) at a locus can result in distorted calculations of the significance in associating SNPs to disease. *Human genetics*. 2018;137():553-567.
[71] G. Hindley, A. A. Shadrin, D. van der Meer, N. Parker, W. Cheng, K. S. O'Connell, S. Bahrami, A. Lin, N. Karadag, B. Holen, et al., Multivariate genetic analysis of personality and cognitive traits reveals abundant pleiotropy. *Nature human behaviour*. 2023;7(9):1584-1600.
[72] D. He, H. Liu, W. Wei, Y. Zhao, Q. Cai, S. Shi, X. Chu, X. Qin, N. Zhang, P. Xu, et al., A longitudinal genome-wide association study of bone mineral density mean and variability in the UK Biobank. *Osteoporosis International*. 2023;34(11):1907-1916.
[73] X. Liu, B. Li, M. W. Vernooij, E. B. Wolvius, G. V. Roshchupkin, and E. E. Bron, AI-based association analysis for medical imaging using latent-space geo metric confounder correction. *Medical Image Analysis*,2025;(102):103529
[74] D. H. Katz, U. A. Tahir, A. G. Bick, A. Pampana, D. Ngo, M. D. Benson, Z. Yu, J. M. Robbins, Z. Chen, D. E. Cruz, et al., Whole genome sequence analysis of the plasma proteome in black adults provides novel insights into cardiovascular disease. *Circulation*. 2022;145(5):357-370.
[75] M. Kirchler, S. Konigorski, M. Norden, C. Meltendorf, M. Kloft, C. Schurmann, C. Lippert. transferGWAS: GWAS of images using deep transfer learning. *Bioinformatics*. 2022;38(14):3621-3628.
[76] Y. Li, Z. Xiong, M. Zhang, P. G. Hysi, Y. Qian, K. Adhikari, J. Weng, S. Wu, S. Du, R. Gonzalez-Jose, et al., Combined genome-wide association study of 136 quantitative ear morphology traits in multiple populations reveal 8 novel loci. *PLoS Genetics*. 2023;19(7):e1010786.
[77] K. Watanabe, E. Taskesen, A. Van Bochoven, D. Posthuma. Functional mapping and annotation of genetic associations with FUMA. *Nature communications*. 2017;8(1):1826.
[78] C. Wijmenga. From LD-based mapping to GWAS. *Nature Reviews Genetics*. 2021;22(8):480-1.
[79] V. Tam, N. Patel, M. Turcott, Y. Bossé, G. Paré, D. Meyre. Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics*. 2019;20(8):467-84.
[80] M. Yuan, S. Goovaerts, M. Vanneste, H. Matthews, H. Hoskens, S. Richmond, O. D. Klein, R. A. Spritz, B. Hallgrimsson, S. Walsh, M.D. Shriver, et al., Mapping genes for human face shape: Exploration of univariate phenotyping strategies. *PLoS computational biology*. 2024;20(12):e1012617.
[81] J. B. Cole, M. Manyama, E. Kimwaga, J. Mathayo, J. R. Larson, D. K. Liberton, K. Lukowiak, T. M. Ferrara, S. L. Riccardi, M. Li, W. Mio, et al., Genomewide association study of African children identifies association of SCHIP1 and PDE8A with facial size and shape. *PLoS genetics*. 2016;12(8):e1006174.
[82] M. Zhang, S. Wu, S. Du, W. Qian, J. Chen, L. Qiao, Y. Yang, J. Tan, Z. Yuan, Q. Peng, Y. Liu, et al., Genetic variants underlying differences in facial morphology in East Asian and European populations. *Nature Genetics*. 2022;54(4):403-11.
[83] F. Liu, F. Van Der Lijn, C. Schurmann, G. Zhu, M. M. Chakravarty, P. G. Hysi, A. Wollstein, O. Lao, M. De Bruijne, M. A. Ikram, A. van der Lugt, et al., A genome-wide association study identifies five loci influencing facial morphology in Europeans. *PLoS genetics*. 2012;8(9):e1002932.

[84] J. K. Pickrell, T. Berisa, J. Z. Liu, L. Ségurel, J. Y. Tung, D. A. Hinds. Detection and interpretation of shared genetic influences on 42 human traits. *Nature genetics*. 2016;48(7):709-17.
[85] Z. Xiong, Y. Li, X. Liu, H. Lu, P. G. Hysi, L. M. Pardo, A. G. Uitterlinden, F. Rivadeneira, M. A. Ikram, M. Ghanbari, E. B. Wolvius, G. V. Roshchupkin, S. Richmond, T. Nijsten, T. D. Spector, S. Wang, F. Liu, M. Kayser. C-GWAS of 946 facial features identifies 188 genetic loci explaining 8 percent of facial variance and enhancing face prediction in modern and archaic humans. *Nature Communication*. 2025;16: 6562