TransferGWAS of T1-weighted Brain MRI Data from the UK Biobank

Alexander Rakowski[1*], Remo Monti[1,2], and Christoph Lippert[1,3]

[1]Digital Health Machine Learning, Hasso Plattner Institute for Digital Engineering, Potsdam, Germany
[2]Max-Delbrück-Center for Molecular Medicine in the Helmholtz Association, Berlin Institute for Medical Systems Biology, Berlin, Germany
[3]Hasso Plattner Institute for Digital Health at Mount Sinai, New York, United States

*Corresponding author
E-mail: alexander.rakowski@hpi.de

Abstract

Genome-wide association studies (GWAS) traditionally analyze single traits, e.g., disease diagnoses or biomarkers. Nowadays, large-scale cohorts such as the UK Biobank (UKB) collect imaging data with sample sizes large enough to perform genetic association testing. Typical approaches to GWAS on high-dimensional modalities extract predefined features from the data, e.g., volumes of regions of interest. This limits the scope of such studies to predefined traits and can ignore novel patterns present in the data. TransferGWAS employs deep neural networks (DNNs) to extract low-dimensional representations of imaging data for GWAS, eliminating the need for predefined biomarkers. Here, we apply transferGWAS on brain MRI data from the UKB. We encoded 36,311 T1-weighted brain magnetic resonance imaging (MRI) scans using DNN models trained on MRI scans from the Alzheimer's Disease Neuroimaging Initiative, and on natural images from the ImageNet dataset, and performed a multivariate GWAS on the resulting features. Furthermore, we fitted polygenic scores (PGS) of the deep features and computed genetic correlations between them and a range of selected phenotypes. We identified 289 independent loci, associated mostly with bone density, brain, or cardiovascular traits, and 14 regions having no previously reported associations. We evaluated the PGS in a multi-PGS setting, improving predictions of several traits. By examining clusters of genetic correlations, we found novel links between diffusion MRI traits and type 2 diabetes.

1 Author Summary

Genome-wide association studies are a popular framework for identifying regions in the genome influencing a trait of interest. At the same time, the growing sample sizes of medical imaging datasets allow for their incorporation into such studies. However, due to high dimensionalities of imaging modalities, association testing cannot be performed directly on the raw data. Instead, one would extract a set of measurements from the images, typically using predefined algorithms, which has several drawbacks - it requires specialized software, which might not be available for new or less popular modalities, and can ignore features in the data, if they have not yet been defined. An alternative approach is to extract the features using pretrained deep neural network models, which are well suited for complex high-dimensional data and have the potential

to uncover patterns not easily discoverable by manual human analysis. Here, we extracted deep feature representations of brain MRI scans from the UK Biobank, and performed a genome-wide association study on them. Besides identifying genetic regions with previously reported associations with brain phenotypes, we found novel regions, as well as ones related to several other traits such as bone mineral density or cardiovascular traits.

2 Introduction

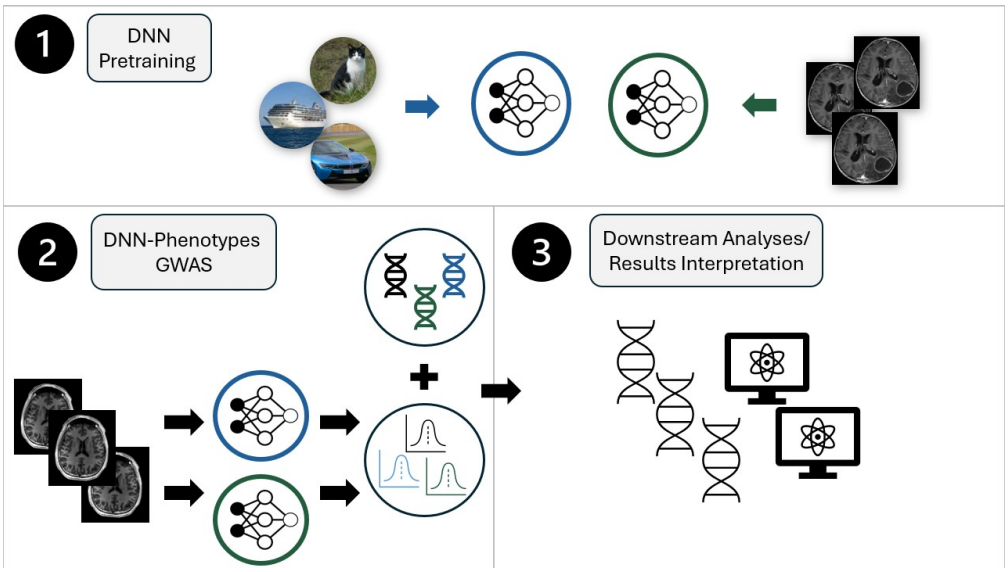
The growing size of medical imaging datasets within biobanks is increasing the power of genome-wide association studies (GWAS) performed on such modalities. For example, the number of associated loci found in a GWAS of phenotypes derived from brain magnetic resonance imaging (MRI) data in the UK Biobank (UKB) increased over 4-fold between the initial study of Elliott et al. [14] and the consecutive study of Smith et al. [49]. The initial approaches to imaging GWAS were based on the extraction of predefined image-derived phenotypes (IDPs) [14, 49, 43]. While being interpretable, such analyses require the availability of automated tools for IDP extraction for the modality of interest and are limited to traits defined a priori, potentially preventing novel genetically-driven phenotypes from being discovered.

Instead of using manually defined traits, a recent line of work employed deep learning (DL) to derive imaging features using pretrained deep neural network (DNN) models to perform GWAS on. This approach has been demonstrated to be successful in a range of imaging modalities, including retinal fundus images [24, 53], cardiovascular magnetic resonance (CMR) images [3, 4], or brain MRI scans [39]. In this work, we perform an imaging GWAS on $N = 36,311$ T1-weighted brain MRI scans from the UKB dataset. As opposed to the ENDO approach of Patel et al. [39], who pretrained a DNN on data from the same dataset where the GWAS was performed, we employed the transferGWAS pipeline of Kirchler et al. [24] and used two DNN models pretrained on other datasets, following a transfer learning methodology. We encoded the brain scans using models pretrained on the ImageNet [48] and Alzheimer's Disease Neuroimaging Initiative (ADNI) datasets [35], with the former extracting "general" image features and the latter focusing on brain MRI and dementia-specific ones. Our GWAS performed on these features identified a number of loci not detected in the IDP or ENDO brain MRI studies, several of which were not reported in any previous GWA studies. We further conducted downstream analyses using the discovered genetic variants, demonstrating their utility in creating more predictive polygenic score (PGS), and pointing to novel genetic correlations between type 2 diabetes (T2D) and diffusion magnetic resonance imaging (dMRI) traits (see Fig 1 for an overview of our workflow).

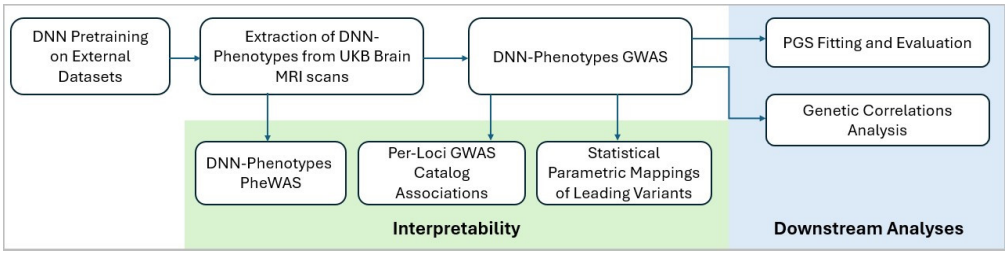
3 Results

3.1 Interpretation of the DNN Features

In order to interpret the signal carried by the DNN features, we extracted the first 10 principal components (PCs) of both DNN models, and performed a phenome-wide association study (PheWAS) against each PC and 7,744 UKB phenotypes (supplementary Table S2). We found 2,408 and 2,622 significantly associated phenotypes for the ImageNet and ADNI PCs respectively, having P-values below the Bonferroni-corrected threshold of $\approx 6.5 \cdot 10^{-7}$. Fig 2 shows the percentage of significantly associated traits per category. The top 35 categories with the highest ratio of significant hits contained 17 brain-related categories, with the other ones being bone density, body composition, or blood-related categories. In almost all cases the ADNI PCs were associated with a higher number of phenotypes than the ImageNet PCs.



(a)



(b)

Figure 1: **Overview of our study and workflow. (a)** A general overview of the study: (1) - we pretrained 2 DNN models on external datasets of natural images, and of brain MRI scans (2) - encoded brain MRI data from the target datasets and performed GWAS on the DNN-derived phenotypes (3) performed a series of downstream analyses using the learned DNN features and discovered genetic variants. **(b)** Description of each step involved in the complete workflow of our study.

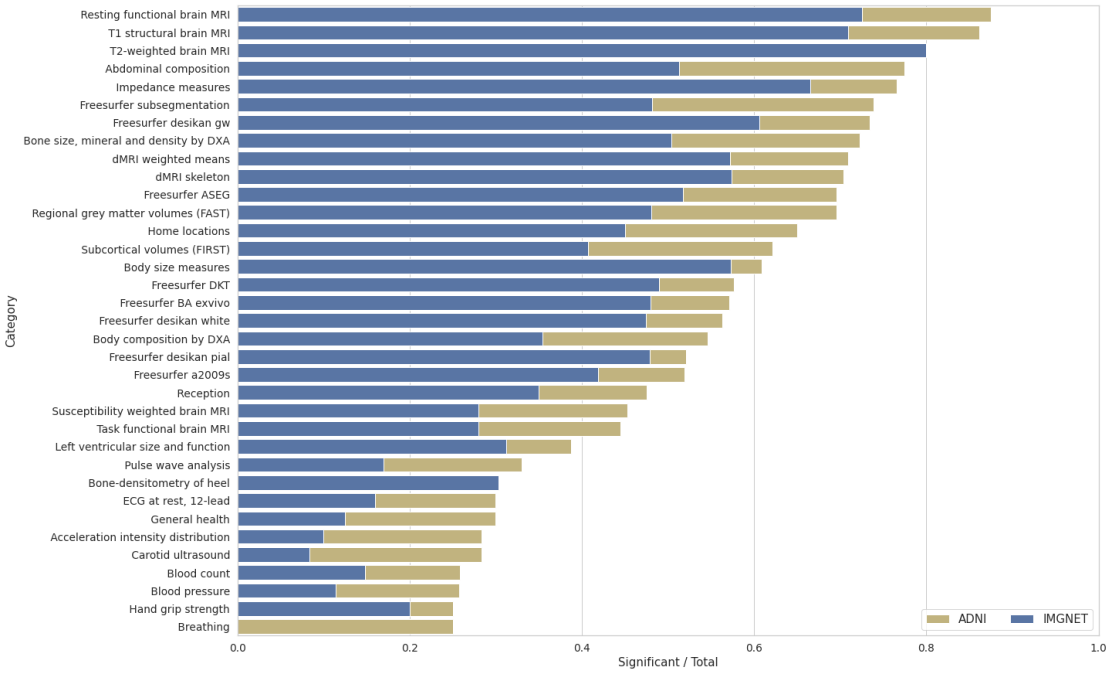


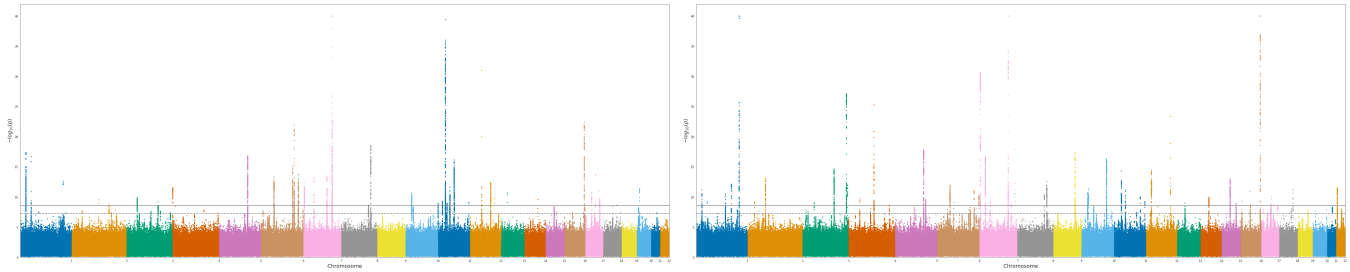
Figure 2: Results of the PheWAS performed on the principal components (PCs) of the ImageNet (blue) and ADNI (yellow) pretrained models. For each phenotype category from the UK Biobank (UKB) we plot the number of significant associations per model divided by the total number of traits in that category. Shown are the top 35 phenotype categories with the highest ratio of significant associations.

3.2 GWAS Results

At the Bonferroni-corrected significance threshold of $2.5 \cdot 10^{-9}$, we found $4,665$ peak associations for the ImageNet and $5,291$ for the ADNI pretrained models, resulting in $4,382$ and $4,360$ distinct variants for ImageNet and ADNI. The clumping procedure then identified 194 and 165 independent regions for the ImageNet and the ADNI models respectively. This amounted to $7,458$ distinct variants and 289 distinct regions across all 20 features of both DNN models. Fig 3 shows the Manhattan plots for both models, aggregated over each of the 10 PCs per model. We estimated the heritability of each PC using linkage disequilibrium score regression (LDSC) (Section 5.5) and found the ADNI-pretrained PCs to be more heritable, with a mean $h^2 = 0.19$, and the ImageNet PCs having a mean $h^2 = 0.13$ (Fig 4). The summary statistics for all PCs are made publicly available as a figshare resource under `https://doi.org/10.6084/m9.figshare.25933717.v1`.

3.2.1 GWAS Catalog Associations

For each independent locus, we queried associations reported in previous GWA studies from the NHGRI-EBI GWAS Catalog [7] (Fig 5). The dominating phenotype categories included bone mineral density (BMD)-related traits and a range of brain traits, such as cortical thickness, diffusion, or volumes of brain regions of interest (ROIs). We note that the ADNI-pretrained features tagged more regions corresponding to brain-related traits, whereas the ImageNet model



(a) ImageNet pretraining.

(b) ADNI pretraining.

Figure 3: Manhattan plots for the GWAS ($n = 36,311$ individuals, $16,472,121$ SNPs) performed on features of the ImageNet (plot a)) and the ADNI (plot b)) pretrained models. The horizontal lines mark the initial significance threshold of $5 \cdot 10^{-8}$ (dashed line) and Bonferroni-corrected threshold of $2.5 \cdot 10^{-9}$ (solid line). For visualization purposes we truncate P-values below 10^{-40} and plot only the minimal P-values across each of the 10 features per model.

tagged more regions related to "general" body structure, such as BMD, height, or body mass index (BMI). Overall, out of the 289 independent loci, 72 did not have brain-related associations reported in the catalog.

Among neuropsychiatric disorders with the highest number of distinct regions, 47 were associated with schizophrenia, 37 with neuroticism, 36 with attention deficit hyperactivity disorder, 35 with bipolar disorder, 33 with depression, 32 with Alzheimer's disease, 30 with autism, 22 with anorexia nervosa and 21 with anxiety.

3 out of the 10 first traits were not directly brain-related: heel bone mineral density (HBMD) (144 regions), total BMD (125 regions), and height (113 regions). The associations between BMD and the brain have been investigated in the context of neurological disorders [29, 30, 54], as well as in samples of healthy subjects [2]. Loskutova et al. [29, 30] reported a correlation between BMD and an early onset of Alzheimer's disease (AD), as well as with several brain volumes. HBMD is postulated to be a causal factor for multiple sclerosis (MS) through an increased risk of fractures [54]. Bae et al. [2] showed that osteoporosis increases the pace of parenchymal atrophy and ventricular enlargement during aging of healthy individuals.

Another prevalent category were blood-related traits, such as cell counts: white (67), red (32), monocyte (45), neutrophil (41), eosinophil (40), lymphocyte, (25) reticulocyte (23), blood pressure (95) or hypertension (45), or hemoglobin (68). Blood pressure and hypertension are known factors influencing brain morphology, as well as cognitive performance or dementia [50, 13, 15, 45], while aenemia is a causal factor for cognitive decline and AD [40, 52].

3.2.2 Novel Loci

In total, we found existing associations for 275 regions in studies conducted on the British population, and 278 regions among all populations. Out of the remaining 11 loci, one was located within an RNA gene, and 10 within 7 distinct protein-coding genes. Among the associated phenotypes, 6 genes were associated with mental or neurodevelopmental disorders such as AD, schizophrenia, or attention-deficit/hyperactivity disorder (ADHD), and 4 genes were associated with T2D.

As a further means of interpreting the novel regions, we computed statistical parametric mappings (SPMs) for each leading single-nucleotide polymorphism (SNP) (Fig 6) and computed the fractions of volume of each brain region correlated with each lead variant (Fig 7). All SNPs

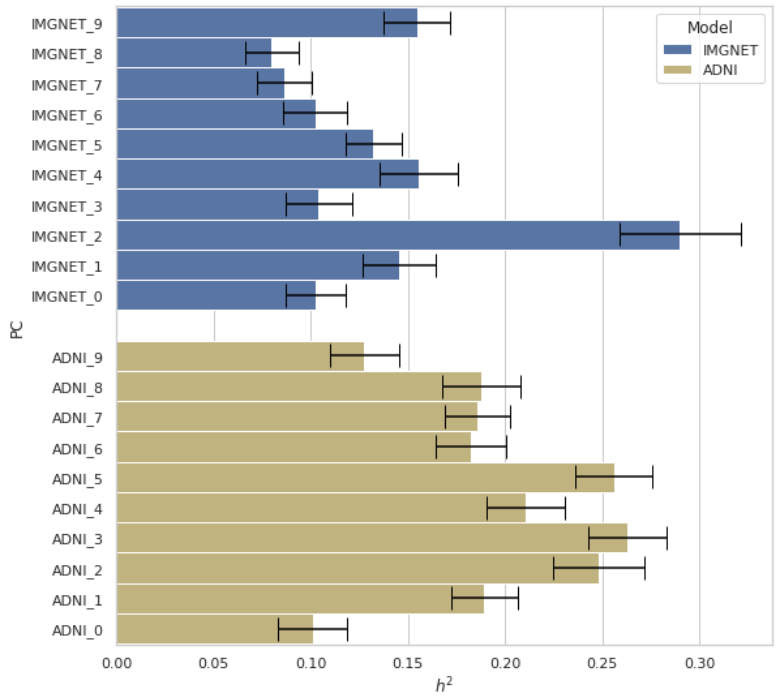


Figure 4: h^2 heritability estimates of principal components (PCs) of the ImageNet (blue) and ADNI (yellow) pretrained models, obtained using linkage disequilibrium score regression (LDSC). Black lines indicate the standard error of the estimates.

were correlated with the left cerebral cortex and cerebrospinal fluid. Most notable was the variant rs111469125 (16:87268090) located inside the C16orf95 gene, being correlated with 19 out of 22 brain regions, in particular with several ventricle structures: the 3rd and 4th ventricles (3% of total voxels), and the left and right lateral ventricles (1% and 1.5% of total voxels). It was also correlated with 3% of the voxels of cerebrospinal fluid, and was the only new variant correlated with the left cerebellum white matter, the left thalamus, and the right caudate.

3.2.3 Comparison with Previous Studies

We performed another GWAS using discovery and replication cohorts (23,604 and 12,709 samples), replicating 1,631 hits over 1,510 unique variants, which amounted to 70 replicated loci. We compared our results with two GWA studies on UKB brain MRI data - the first one using 3,144 brain imaging-derived phenotypes [49] and the second study using 256 DL-based features [39], which yielded 692 and 43 replicated loci respectively. Out of our 70 replicated loci, 9 were not present in the 692 of Smith et al. [49], and 28 were not present in the 43 loci of Patel et al. [39].

3.3 TransferGWAS Polygenic Scores

Here, we evaluated the potential of variants discovered in our study for downstream prediction of phenotypes using 20 PGS fitted for each of the 20 DNN PCs with the summary statistics from our GWAS. In order to compute the features of the DNN models, imaging data need to be present, which constitutes less than a tenth of all UKB samples. On the other hand, genotyping

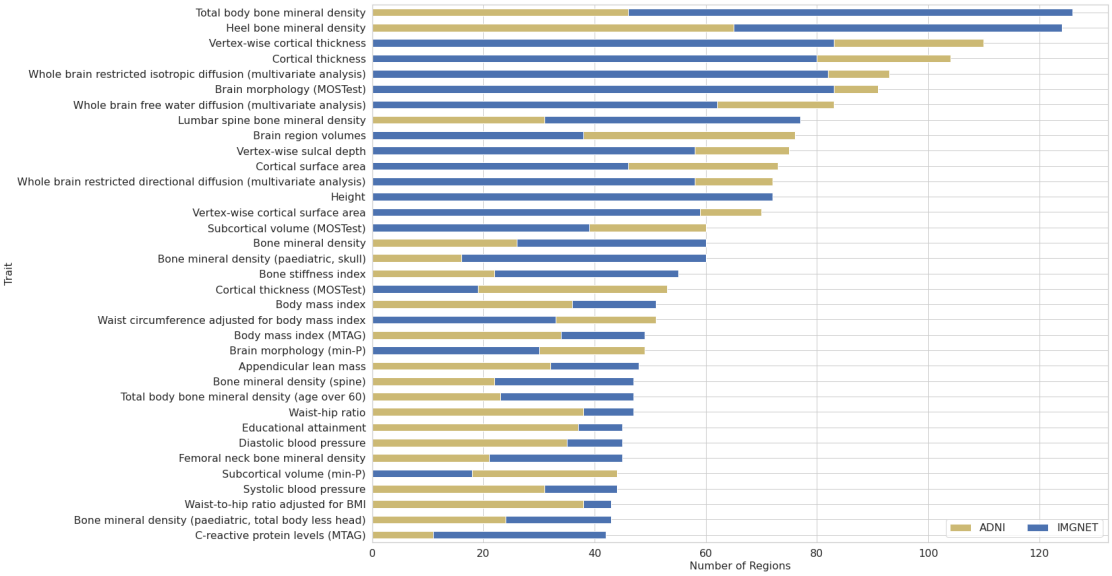


Figure 5: Number of independent loci per trait with associations reported in previous studies included in the NHGRI-EBI GWAS Catalog [7]. Shown are the top 35 traits with the highest number of associated regions.

data were available for all participants. This allowed us to calculate the PGS for all remaining $N = 451,450$ participants not included in the GWAS sample. The corresponding methods are described in Section 5.4, while the weights of the fitted scores are made publicly available as a figshare resource under `https://doi.org/10.6084/m9.figshare.25933663.v1`.

3.3.1 PGS PheWAS

To gain insights into which traits the PGS might be predictive of, we performed a PheWAS on the 20 PGS and the 7,744 UKB phenotypes (supplementary Table S2). Note that while the "raw" DNN PCs can encode both genetic and environmental signals, the PGS should capture only the former, and thus we expected the associations between the phenotypes to differ from the PheWAS performed on the PCs. The total number of significant PC-phenotype associations and the effect sizes were higher for the original PCs than for the PGS: 28,767 vs. 25,948 significant associations in total, 2,928 vs. 2,860 distinct associated traits, with mean effect sizes of $\bar{\beta} = 0.08$ vs. $\bar{\beta} = 0.04$. We identified 3 potentially interesting groups of associations (Fig 8):

- traits related to BMD
- weight/fat mass/BMI
- cardiovascular traits

which we decided to investigate further in a prediction setting.

3.3.2 Predictive Performance Compared to Trait-Specific PGS

We tested the utility of our developed PGS by evaluating whether they can improve predictions of phenotypes from UKB over PGS designed specifically for particular traits in a multi-PGS

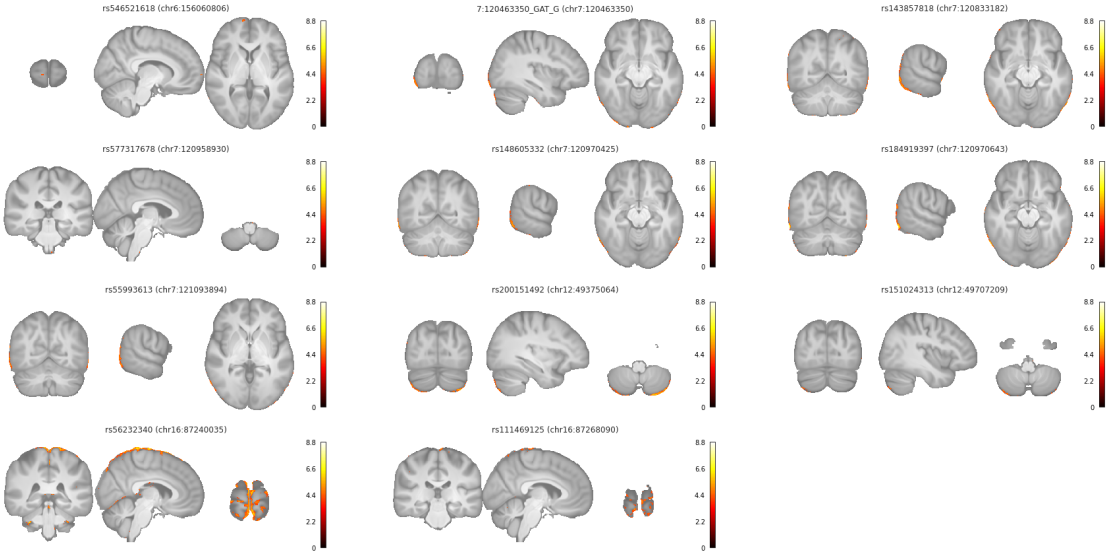


Figure 6: Statistical parametric mappings (SPMs) for genetic regions with no previously reported GWAS associations. Plotted are values of the t-statistics of the correlation coefficients between lead variants of each region and each single voxel in the MRI scans. We plot values below the Bonferroni-corrected significance threshold accounting for the total number of voxels tested.

setting [25]. We chose a set of 9 phenotypes based on the PheWAS results and computed their corresponding scores using PGS available in the PGS Catalog [26]. For each phenotype, we then fitted and evaluated two linear models: one fitted using only the trait-specific PGS, and one additionally using our transferGWAS PGS. While there were statistically significant improvements in predictions for 4 out of 9 traits, they yielded arguably small performance increases ($\sim 1.5\%$ of relative improvement), with the exception of predicting HBMD using a (general) BMD PGS, where the relative improvement was over 20% (Table 1). We decided to further investigate the HBMD results. Since the improvement could have been stemming from a lower signal in the dataset of the external PGS, compared to the UKB, we conducted a further comparison within the UKB. We performed a GWAS on HBMD using $19,909$ samples from our GWAS data which had HBMD measurements available, and $16,404$ randomly drawn from the remaining samples of the white-british participants, to match the sample size and population of our GWAS, and fitted a new PGS on the resulting summary statistics. We then evaluated the newly-created HBMD PGS with and without our transferGWAS PGS on the remaining UKB data and observed the same relative improvement of 1% in performance ($p < 0.001$). This indicates that transferGWAS has the potential to identify additional variants for related traits while using the sample size. We hypothesize that this might be due to certain pleiotropic variants having a larger effect on the DNN PCs than on HBMD, and thus being able to be detected with our DL GWAS and not with the HBMD-dedicated GWAS.

3.4 Genetic Correlations

The results of the PheWAS conducted on the learned PCs led us to a set of traits that we decided to investigate further. In order to analyze the genetic components of the PCs, we computed genetic correlation coefficients between 102 selected traits and each of the 20 PCs

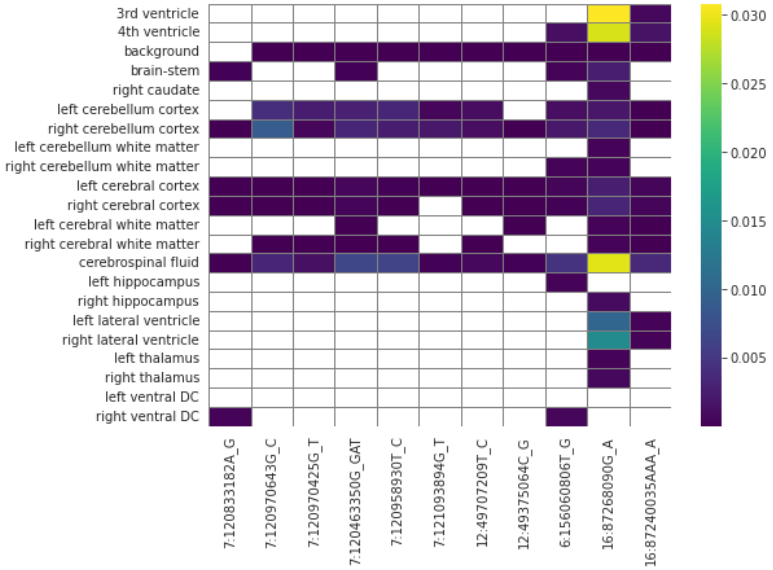


Figure 7: Fractions of volume of brain regions correlated with genetic regions with no previously reported GWAS associations. The values are computed as the total number of voxels in a given brain region significantly correlated with a lead variant, divided by the total number of voxels in that brain region. White cells indicate no voxels being significantly correlated for a given brain region-genetic region pair.

(see Section 5.5 for details). 39 traits were significantly correlated, surpassing the Bonferroni-corrected threshold of $\approx 2.5 \cdot 10^{-5}$. We grouped the traits into 3 groups:

- (volumes of) brain ROIs (e.g., ventricles, brain stem, cerebrospinal fluid (CSF))
- dMRI traits (e.g., fractional anisotropy (FA), orientation dispersion index (ODI))
- "general" traits: Height, T2D, BMI, HBMD

Additionally, we tested for correlations with AD, educational attainment, and unipolar depression, finding no significant correlations when corrected for multiple testing ($p > 0.001$). The significantly associated traits are shown in Fig 9, where we observed several "clusters" of PC-trait associations.

Several PCs were associated with volumes of multiple brain ROIs. The first two PCs of ImageNet (IMGNET0, IMGNET1) seemed to capture the overall body size, as they were negatively correlated with height and white matter volume, and positively with ventricular ROIs and CSF.

PCs ADNI2, ADNI3, and IMGNET4 were genetically associated with volumes of several brain ROIs, e.g., cerebral white matter, putamen, or thalamus. ADNI2 and ADNI3 were also associated with volumes of CSF and the lateral ventricle. Interestingly, ADNI2 had a positive correlation both for CSF and the lateral ventricle, as well as for gray and white matter structures, whereas one might expect the ventricular volumes (and thus CSF) to grow with the shrinkage of brain structures.

PCs associated with HBMD seemed to capture different aspects of brain anatomy. IMGNET2 had a negative correlation with HBMD, BMI, and cerebral white matter, but also with multiple ventricular volumes. On the other hand, ADNI8 and IMGNET4 also had negative genetic correlations with HBMD, but positive ones with cerebral white matter.

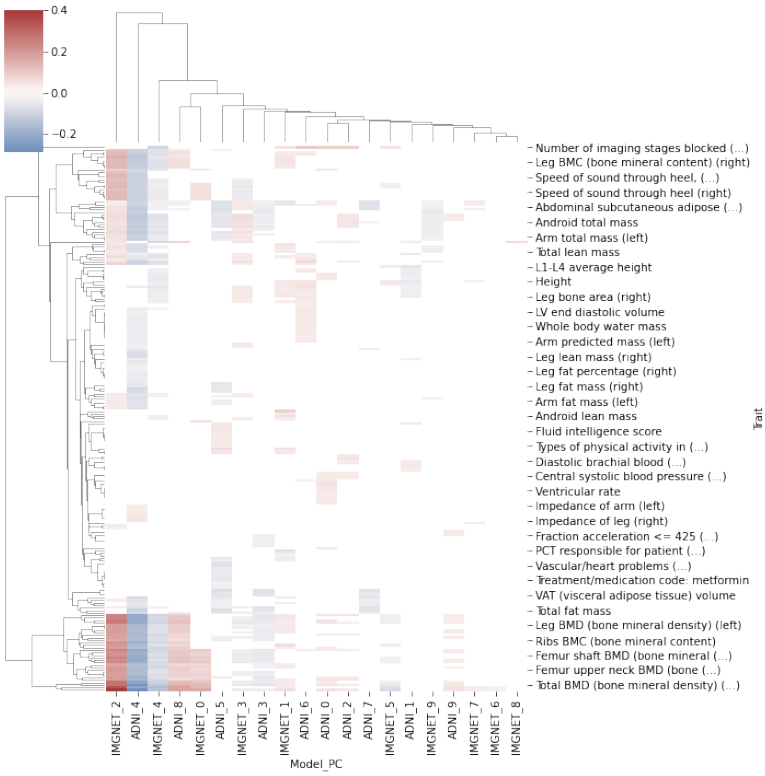


Figure 8: Results of the PheWAS ($N = 451,450$) between the polygenic scores (PGS) fitted on the features of the DL models (rows) and phenotypes from the UK Biobank (UKB) dataset (columns). Cell colors represent the magnitudes and the signs of the estimated association coefficients between each PGS and phenotype combination.

ADNI0 and ADNI4 were associated with a range of Diffusion MRI traits, as well as with several ventricular ROIs. Furthermore, ADNI4 was genetically correlated with HBMD and BMI, and was the only PC associated with T2D, which we further discuss below.

3.4.1 ADNI4 and T2D

BMI was shown to increase the risk of developing T2D [8, 17], as well as being genetically correlated to T2D [8]. The signs of genetic correlations between ADNI4, and BMI and T2D were also matching. ADNI4 was also positively genetically associated with HBMD. T2D patients have been shown to have a higher bone density [42, 27]. Evidence also exists for shared heritability between BMD and T2D, albeit relatively small [47, 55]. As with BMI, the sign of the genetic correlation between ADNI4 and HBMD was positive. Regarding the brain ROIs, ADNI4 was positively correlated with volumes of the lateral, 3rd, and 4th ventricles, as well as with the CSF. Ventricular enlargement and increase in CSF are associated with several neurodegenerative diseases, such as AD, MS, or schizophrenia [9, 36]. Several studies showed an association between T2D and volumes of white matter structures (whole brain volume, frontal lobe), gray matter (overall trend in all structures), as well as CSF and ventricular volumes [33]. Furthermore, ADNI4 was genetically correlated with 35 different dMRI traits:

Trait	PGS Catalog	Our PGS + PGS Catalog	Δ
Height	0.759	0.759	0.000
BMI	0.284	0.284	0.000
Heel bone mineral density (1)	0.273	0.275	0.002*
Heel bone mineral density (2)	0.064	0.078	0.014*
Red blood cell count	0.387	0.387	0.000
White blood cell count	0.104	0.106	0.002*
Systolic blood pressure	0.283	0.284	0.000
Diastolic blood pressure	0.175	0.176	0.001*
Ventricular rate	0.027	0.025	-0.001

Table 1: Comparison of predictive performance of Multi-PGS models using only trait-specific polygenic scores (PGS) (2nd column) and including our TransferGWAS PGS (3rd column) for a set of selected phenotypes from UK Biobank (UKB), measured with the R^2 coefficient of determination. Significant differences are marked with (*). Heel bone mineral density (1) and (2) correspond to results of using PGS for heel bone mineral density, or (general) bone mineral density respectively.

Mean diffusivity (MD) traits 4 MD traits were positively genetically correlated with ADNI4: fornix, superior serebellar peduncle (both sides), and the superior fronto-occipital fasciculus (left). Positive associations between T2D and MD have been found in observational studies [19, 46].

Fractional anisotropy (FA) traits FA traits have been found to be negatively correlated with T2D in literature [19, 46, 33]. We found 4 traits to be negatively genetically correlated with ADNI4, however the posterior limb of left internal capsule was positively genetically correlated with the PC. The direction of this correlation seemed to be in opposition to the associations found in observational studies [33]. On the other hand, it is postulated to be causal with the same sign for fasting insulin [12], an increase of which is an indicator of T2D. We identified two regions containing shared variants located at Chr2:27766284 and Chr14:91881387. The first region contains missense and intron variants for GCKR gene (ENSG00000133962), a glucokinase regulator, with no previously reported associations for brain phenotypes, missense and intron variants for C2orf16 (ENSG00000221843) and intron variants for ZNF512 (ENSG00000243943) both protein coding genes with association with neurodegenerative diseases, T2D, and blood measurements. The second region contains intron variants for the CCDC88C (ENSG00000015133), a protein coding gene, with associations with glucose metabolism, brain measurements, and neurodegenerative diseases, and CCDC88C-DT (ENSG00000258798), and RNA gene that is a divergent transcript for CCDC88C, with associations with brain measurements and hypertension. The above may be another indicator of a non-trivial relation between FA of limb of internal capsule and T2D, with potentially different shared heritability and environmental effects.

Orientation dispersion index (ODI) traits 7 ODI traits were positively correlated with ADNI4, while 3 traits were correlated negatively. ODI of white matter tracts was reported to be positively correlated with duration of T2D and with levels of HbA1c, a marker for T2D, while ODI of internal capsule was reported to have a negative correlation [1], which is consistent with 9 out of 10 of our findings. We found a negative genetic correlation for the posterior right corona radiata, which had shared variants in regions located at chr8:119486034 and 11:27465591. The first region has intron variants for SAMD12 (ENSG00000177570), a protein coding gene

with associations with brain measurements, MS, bone density, and blood measurements. The second region has intron variants for LGR4 (ENSG00000205213), a protein coding gene with associations with brain measurements, bone density, and body mass traits.

Mode of anisotropy (MO) traits 5 MO traits were genetically positively correlated with ADNI4, and 6 negatively. Fasting insulin, a marker for T2D was reported to be negatively associated with anterior corona radiata [12]. We found positive genetic correlations between both sides of the posterior and superior corona radiata and ADNI4, with shared variants with T2D located in the region chr2:27766284 for the superior, and in chr8:119486034 for the posterior. The first region contains missense and intron variants for the GCKR and C2orf16 genes, and an intron variant for ZNF512 (see the FA regions), and the second region has intron variants for SAMD12 (see the ODI regions above). The correlations between the other 7 traits are reported in supplementary Table S3.

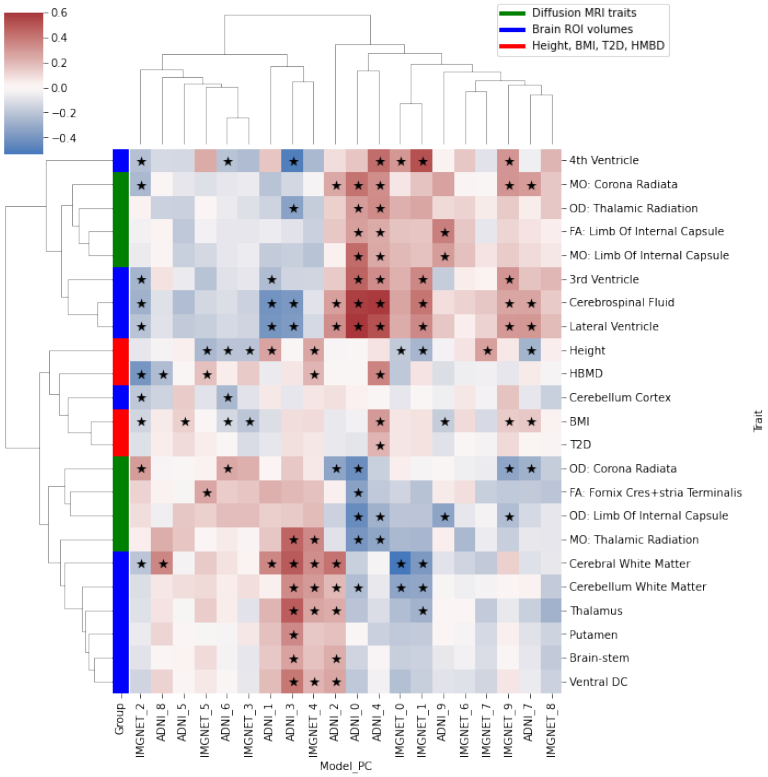


Figure 9: Genetic correlation coefficients between the 20 deep neural network (DNN) principal components (PCs) (rows) and 23 significantly associated phenotypes (columns), out of 27 candidate traits from the UK Biobank (UKB). Cell colors represent the magnitudes and the signs of the estimated genetic correlation coefficients between each PC and phenotype combination.

4 Discussion

Using the transferGWAS approach we performed a GWAS on 20 DNN feature representations of 36,311 T1-weighted brain MRI scans from the UKB, identifying 289 loci, 11 of them without any previously reported associations, and 72 without any associations for brain-related traits. Similar to the findings of the initial transferGWAS study of retinal fundus images of Kirchler et al. [24], the features of an ImageNet-pretrained model were associated with a higher number of loci related to "general" body structure traits, such as BMD or BMI, whereas features from a model pretrained directly on brain MRI data identified more loci corresponding to brain measurements and neurodegenerative diseases. Overall, features of both DNN models were associated directly, through PheWAS, or genetically, through GWAS-identified loci, with a large number of BMD traits. For example, the ImageNet and ADNI-derived features were significantly associated with over 50% and 70% of phenotypes under the UKB category 125 "Bone size, mineral, and density by DXA", and with over 120 and 40 distinct loci associated "Total body bone mineral density" in the NHGRI-EBI GWAS Catalog. Detecting these genetic regions in features derived from brain MRI data seems to confirm the connections between BMD and brain measurements, as well as with neurodegenerative diseases previously reported in the literature (as discussed in Section 3.2.1), which we further investigated with an analysis of genetic correlations (Section 3.4), highlighting particular brain ROIs genetically associated with BMD. Furthermore, the genetic correlations identified by our study shed more light on the relations between dMRI measurements and T2D, BMI, as well as cardiovascular traits, also reported in several studies (Section 3.4.1). Finally, we demonstrated a practical application of our findings by constructing PGS of our DNN-derived phenotypes, which improved predictions of existing PGS of BMD, white cell blood count, or diastolic blood pressure. In a further analysis, we fitted a PGS directly to HBMD measurements on a UKB sample of the same size as our GWAS and observed the same improvement in performance when augmented with our DNN PGS, indicating that the transferGWAS approach can identify additional variants for a trait of interest, being complementary to conducting a trait-dedicated GWAS.

We demonstrated how transferGWAS can be applied to discover new variants and in turn, lead to better phenotype predictions. However, a drawback of using features of pretrained DNN models as traits of interest is their reduced interpretability compared to predefined phenotypes. While we analyzed both the DNN-derived traits and the discovered loci with a range of techniques (PheWAS, querying the GWAS Catalog, SPMs), we highlight the need for further developing apossibly automated pipeline for interpretability of the DNN features, to foster their utility for consecutive research and clinical applications.

5 Materials and Methods

5.1 Pretraining of the Neural Network Models

The first model used for feature extraction was trained on 4,480 T1-weighted scans from the ADNI dataset [41]. The network architecture was a 3D convolutional variational autoencoder (VAE) [23], trained in a multi-task manner. The model consisted of 3 sub-networks: an encoder, a decoder, and a prediction head. The 256-dimensional outputs of the encoder network constituted the latent representations of the input data. The first task was the standard VAE objective, i.e., reconstructing the input scans from the latent representations, while regularizing the representations to match a standard normal prior distribution with a Kullback-Leibler divergence (KLD) loss term. The second task was to predict the clinical dementia rating (CDR) from the latent representations. The aim of the VAE objective was to learn general structural features

describing an MRI scan, while the prediction task should promote neurodegenerative features associated with the presence of dementia. Additionally, we input the age and sex of each participant into the decoder and prediction networks, forcing the model to learn latent representations invariant to age and sex, and thus potentially increasing the statistical power of the GWAS. We trained the model for 500 epochs with the Adam optimizer [22], with a mini-batch size of 128. The weights of the reconstruction, KLD, and the predictions loss terms were 1, 10^{-5} and 10^{-2} respectively. For data preprocessing, we skull-stripped each scan using using the HD-BET tool [20], performed a non-linear registration to the MNI152 template with a 1mm^3 resolution using the FLIRT and FNIRT commands from the FSL software [21], and finally downsampled the scans to a size of $96 \times 96 \times 96$ voxels each.

Following Kirchler et al. [24], we also employed a 2D ResNet50 [18] model pretrained on ImageNet, a non-medical dataset of natural images [48]. We used a readily available trained model from the PyTorch library [38]. We selected the 2048-dimensional output of the penultimate layer as the latent features used for the GWAS. Since the model was trained on 2D data, we could not directly extract features from the 3D MRI scans. Instead, for each scan, we computed the features over each single slice across the axial axis and averaged the results into a single vector.

5.2 GWAS

We selected a sample of $N = 36,311$ UKB participants who "self-identified as 'White British' and have very similar genetic ancestry based on a principal components analysis of the genotypes" (UKB field 22006). We performed the association testing within the linear mixed model (LMM) framework using the BOLT-LMM software [28]. We adjusted for confounding using age, sex, the identifiers of the genotyping array and UKB assessment center, and the first 10 genetic principal components. We filtered the SNPs with the following criteria: MAF$\geq$ 0.1%, Hardy-Weinberg Equilibrium with a significance level of 0.001, pairwise LD-pruning with $R^2 = 0.8$, and maximum missingness of 10% per SNP and participant, which resulted in 577, 570 directly genotyped SNPs. Including imputed genotype data resulted in 16,472,121 variants in total, on which we performed the GWAS. We clumped the variants into independent loci using the PLINK software [44], with a physical distance threshold of 250kb and a significance threshold of 10^{-9} for the index SNPs. We queried the NHGRI-EBI GWAS Catalog [7] using the LDtrait web application [31], with an R^2 cutoff of 10^{-1} and a 250kb window.

5.3 PheWAS

We performed the PheWAS on the PCs of both pretrained models using the PHESANT software [32], with a P-value threshold of $\approx 6.5 \cdot 10^{-7}$ from the Bonferroni correction to account for 20 PCs and 7, 744 different phenotypes from UKB, adjusting for age and sex.

5.4 Polygenic Scores

We fitted the DNN PGS and the custom HBMD PGS using the PRScs method [16], with the prspipe software [37, 34]. For the predictive performance comparison, we queried the PGS Catalog [26] API for a list of PGS developed for each of the 9 phenotypes, ignoring scores that used the UKB for development, to avoid data leakage. We then computed scores for the $N = 451,450$ participants who were not in our GWAS sample using the PGS Catalog Calculator [51]. For each phenotype, we fitted a baseline linear model using all corresponding trait-specific PGS and covariates (age, sex, UKB assessment center, UKB genotyping batch, all UKB genetic PCs)

and another linear model which additionally included our 20 DNN PGS. We used 60% of the data for model fitting and evaluated it on the remaining 40%. We computed P-values for differences between achieved R^2 scores of the two linear models using permutation tests with $1,000$ permutations, randomly selecting predictions from either model for each test sample in each permutation.

5.5 Genetic Correlations

To compute the genetic correlation scores between the PCs and selected traits, we used the LDSC method [6, 5]. We used the provided LD scores precomputed on 1000 Genomes data [10] over HapMap3 [11] SNPs, and used the default values for other parameters of the LDSC. In order to find regions potentially contributing to the genetic correlations between ADNI4, T2D, and dMRI traits (Section 3.4.1), we selected SNPs with a P-value below 0.0001 for which the magnitude of the product of the z-scores between both ADNI4 and T2D, and ADNI4 and a dMRI trait exceeded a threshold of 15. For the dMRI traits, we selected pairs where the sign of the product of the z-scores matched the sign of the genetic correlation with ADNI4. We consider a region a set of variants within $250,000$ base pairs from a "central" variant.

6 Acknowledgments

This research has received funding by the European Commission in the Horizon 2020 project INTERVENE (Grant agreement ID: 101016775), and was partially funded by the HPI Research School on Data Science and Engineering.

References

[1] Abdulmajeed Alotaibi, Anna Podlasek, Amjad AlTokhis, Chris R Tench, Ali-Reza Mohammadi-Nejad, Stamatios N Sotiropoulos, Cris S Constantinescu, Sieun Lee, and Rob A Dineen. White matter microstructural alteration in type 2 diabetes: A combined uk biobank study of diffusion tensor imaging and neurite orientation dispersion and density imaging. *medRxiv*, pages 2022–07, 2022.

[2] In-Suk Bae, Jae Min Kim, Jin Hwan Cheong, Je Il Ryu, and Myung-Hoon Han. Association between bone mineral density and brain parenchymal atrophy and ventricular enlargement in healthy individuals. *Aging (albany NY)*, 11(19):8217, 2019.

[3] Rodrigo Bonazzola, Nishant Ravikumar, Rahman Attar, Enzo Ferrante, Tanveer Syeda-Mahmood, and Alejandro F Frangi. Image-derived phenotype extraction for genetic discovery via unsupervised deep learning in cmr images. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part V 24*, pages 699–708. Springer, 2021.

[4] Rodrigo Bonazzola, Enzo Ferrante, Nishant Ravikumar, Yan Xia, Bernard Keavney, Sven Plein, Tanveer Syeda-Mahmood, and Alejandro F Frangi. Unsupervised ensemble-based phenotyping helps enhance the discoverability of genes related to heart morphology. *arXiv preprint arXiv:2301.02916*, 2023.

[5] Brendan Bulik-Sullivan, Hilary K Finucane, Verneri Anttila, Alexander Gusev, Felix R Day, Po-Ru Loh, ReproGen Consortium, Psychiatric Genomics Consortium, Genetic Consortium

for Anorexia Nervosa of the Wellcome Trust Case Control Consortium 3, Laramie Duncan, et al. An atlas of genetic correlations across human diseases and traits. *Nature genetics*, 47 (11):1236–1241, 2015.

[6] Brendan K Bulik-Sullivan, Po-Ru Loh, Hilary K Finucane, Stephan Ripke, Jian Yang, Schizophrenia Working Group of the Psychiatric Genomics Consortium, Nick Patterson, Mark J Daly, Alkes L Price, and Benjamin M Neale. Ld score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature genetics*, 47(3): 291–295, 2015.

[7] Annalisa Buniello, Jacqueline A. L. MacArthur, Maria Cerezo, Laura W. Harris, James D. Hayhurst, Cinzia Malangone, Aoife McMahon, Joannella Morales, Edward Mountjoy, Elliot Sollis, Daniel Suveges, Olga Vrousgou, Patricia L. Whetzel, M. Ridwan Amode, Jose A. Guillen, Harpreet Singh Riat, Stephen J. Trevanion, Peggy Hall, Heather Junkins, Paul Flicek, Tony Burdett, Lucia A. Hindorff, Fiona Cunningham, and Helen E. Parkinson. The nhgri-ebi gwas catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Research*, 47(Database-Issue):D1005–D1012, 2019. URL `http://dblp.uni-trier.de/db/journals/nar/nar47.html#BunielloMCHHMMM19`.

[8] S Carlsson, A Ahlbom, P Lichtenstein, and T Andersson. Shared genetic influence of bmi, physical activity and type 2 diabetes: a twin study. *Diabetologia*, 56:1031–1035, 2013.

[9] Adam Conley. *Ventricular Enlargement*, pages 2600–2601. Springer New York, New York, NY, 2011. ISBN 978-0-387-79948-3. doi: 10.1007/978-0-387-79948-3_375. URL `https://doi.org/10.1007/978-0-387-79948-3_375`.

[10] 1000 Genomes Project Consortium et al. A global reference for human genetic variation. *Nature*, 526(7571):68, 2015.

[11] International HapMap 3 Consortium et al. Integrating common and rare genetic variation in diverse human populations. *Nature*, 467(7311):52, 2010.

[12] Shan-Shan Dong, Ke Yu, Shi-Hao Tang, Jing Guo, Yan Guo, and Tie-Lin Yang. Evaluation of the causal associations between brain imaging-derived phenotypes and type 2 diabetes: a bidirectional mendelian randomization study. *medRxiv*, pages 2022–12, 2022.

[13] C Dufouil, A de Kersaint-Gilly, V Besancon, C Levy, E Auffray, L Brunnereau, A Alperovitch, and C Tzourio. Longitudinal study of blood pressure and white matter hyperintensities: the eva mri cohort. *Neurology*, 56(7):921–926, 2001.

[14] Lloyd T Elliott, Kevin Sharp, Fidel Alfaro-Almagro, Sinan Shi, Karla L Miller, Gwenaëlle Douaud, Jonathan Marchini, and Stephen M Smith. Genome-wide association studies of brain imaging phenotypes in uk biobank. *Nature*, 562(7726):210–216, 2018.

[15] Michael J Firbank, Rebecca M Wiseman, Emma J Burton, Brian K Saxby, John T O'Brien, and Gary A Ford. Brain atrophy and white matter hyperintensity change in older adults and relationship to blood pressure: brain atrophy, wmh change and blood pressure. *Journal of neurology*, 254:713–721, 2007.

[16] Tian Ge, Chia-Yen Chen, Yang Ni, Yen-Chen Anne Feng, and Jordan W Smoller. Polygenic prediction via bayesian regression and continuous shrinkage priors. *Nature communications*, 10(1):1776, 2019.

[17] Natallia Gray, Gabriel Picone, Frank Sloan, and Arseniy Yashkin. The relationship between bmi and onset of diabetes mellitus and its complications. *Southern medical journal*, 108(1): 29, 2015.

[18] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[19] Jung-Lung Hsu, Yen-Ling Chen, Jyu-Gang Leu, Fu-Shan Jaw, Cheng-Hui Lee, Yuh-Feng Tsai, Chien-Yeh Hsu, Chyi-Huey Bai, and Alexander Leemans. Microstructural white matter abnormalities in type 2 diabetes mellitus: a diffusion tensor imaging study. *Neuroimage*, 59 (2):1098–1105, 2012.

[20] Fabian Isensee, Marianne Schell, Irada Pflueger, Gianluca Brugnara, David Bonekamp, Ulf Neuberger, Antje Wick, Heinz-Peter Schlemmer, Sabine Heiland, Wolfgang Wick, et al. Automated brain extraction of multisequence mri using artificial neural networks. *Human brain mapping*, 40(17):4952–4964, 2019.

[21] Mark Jenkinson, Christian F Beckmann, Timothy EJ Behrens, Mark W Woolrich, and Stephen M Smith. Fsl. *Neuroimage*, 62(2):782–790, 2012.

[22] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2014. URL `https://arxiv.org/abs/1412.6980`.

[23] Diederik P Kingma and Max Welling. Auto-encoding variational bayes, 2013. URL `https://arxiv.org/abs/1312.6114`.

[24] Matthias Kirchler, Stefan Konigorski, Matthias Norden, Christian Meltendorf, Marius Kloft, Claudia Schurmann, and Christoph Lippert. transferGWAS: GWAS of images using deep transfer learning. *Bioinformatics*, 38(14):3621–3628, 05 2022. ISSN 1367-4803. doi: 10.1093/bioinformatics/btac369. URL `https://doi.org/10.1093/bioinformatics/btac369`.

[25] Eva Krapohl, Hamel Patel, Stephen Newhouse, Charles J Curtis, Sophie von Stumm, Philip S Dale, Delilah Zabaneh, Gerome Breen, Paul F O'Reilly, and Robert Plomin. Multi-polygenic score approach to trait prediction. *Molecular psychiatry*, 23(5):1368–1374, 2018.

[26] Samuel A Lambert, Laurent Gil, Simon Jupp, Scott C Ritchie, Yu Xu, Annalisa Buniello, Aoife McMahon, Gad Abraham, Michael Chapman, Helen Parkinson, et al. The polygenic score catalog as an open database for reproducibility and systematic evaluation. *Nature Genetics*, 53(4):420–425, 2021.

[27] Bo Liu, Jingshuang Liu, Junpeng Pan, Chengliang Zhao, Zhijie Wang, and Qiang Zhang. The association of diabetes status and bone mineral density among us adults: evidence from nhanes 2005–2018. *BMC endocrine disorders*, 23(1):1–9, 2023.

[28] Po-Ru Loh, George Tucker, Brendan K Bulik-Sullivan, Bjarni J Vilhjalmsson, Hilary K Finucane, Rany M Salem, Daniel I Chasman, Paul M Ridker, Benjamin M Neale, Bonnie Berger, et al. Efficient bayesian mixed-model analysis increases association power in large cohorts. *Nature genetics*, 47(3):284–290, 2015.

[29] Natalia Loskutova, Robyn A Honea, Eric D Vidoni, William M Brooks, and Jeffrey M Burns. Bone density and brain atrophy in early alzheimer's disease. *Journal of Alzheimer's Disease*, 18(4):777–785, 2009.

[30] Natalia Loskutova, Robyn A Honea, William M Brooks, and Jeffrey M Burns. Reduced limbic and hypothalamic volumes correlate with bone density in early alzheimer's disease. *Journal of Alzheimer's Disease*, 20(1):313–322, 2010.

[31] Mitchell J Machiela and Stephen J Chanock. Ldlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics*, 31(21):3555–3557, 2015.

[32] Louise AC Millard, Neil M Davies, Tom R Gaunt, George Davey Smith, and Kate Tilling. Software Application Profile: PHESANT: a tool for performing automated phenome scans in UK Biobank. *International Journal of Epidemiology*, 47(1):29–35, 10 2017. ISSN 0300-5771. doi: 10.1093/ije/dyx204. URL `https://doi.org/10.1093/ije/dyx204`.

[33] Hossein Sanjari Moghaddam, Farzaneh Ghazi Sherbaf, and Mohammad Hadi Aarabi. Brain microstructural abnormalities in type 2 diabetes mellitus: A systematic review of diffusion tensor imaging studies. *Frontiers in Neuroendocrinology*, 55:100782, 2019.

[34] Remo Monti, Lisa Eick, Georgi Hudjashov, Kristi Läll, Stavroula Kanoni, Brooke N Wolford, Benjamin Wingfield, Oliver Pain, Sophie Wharrie, Bradley Jermy, et al. Evaluation of polygenic scoring methods in five biobanks reveals greater variability between biobanks than between methods and highlights benefits of ensemble learning. *medRxiv*, pages 2023–11, 2023.

[35] Susanne G Mueller, Michael W Weiner, Leon J Thal, Ronald C Petersen, Clifford Jack, William Jagust, John Q Trojanowski, Arthur W Toga, and Laurel Beckett. The alzheimer's disease neuroimaging initiative. *Neuroimaging Clinics*, 15(4):869–877, 2005.

[36] Camila Orellana, Daniel Ferreira, J Muehlboeck, Patrizia Mecocci, Bruno Vellas, Magda Tsolaki, Iwona Kłoszewska, Hilkka Soininen, Simon Lovestone, Andrew Simmons, et al. Measuring global brain atrophy with the brain volume/cerebrospinal fluid index: normative values, cut-offs and clinical associations. *Neurodegenerative Diseases*, 16(1-2):77–86, 2016.

[37] Oliver Pain, Kylie P Glanville, Saskia P Hagenaars, Saskia Selzam, Anna E Fürtjes, Héléna A Gaspar, Jonathan RI Coleman, Kaili Rimfeld, Gerome Breen, Robert Plomin, et al. Evaluation of polygenic prediction methodology within a reference-standardized framework. *PLoS genetics*, 17(5):e1009021, 2021.

[38] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019.

[39] Khush Patel, Ziqian Xie, Hao Yuan, Sheikh Muhammad Saiful Islam, Wanheng Zhang, Assaf Gottlieb, Han Chen, Luca Giancardo, Alexander Knaack, Evan Fletcher, et al. New phenotype discovery method by unsupervised deep representation learning empowers genetic association studies of brain imaging. *medRxiv*, pages 2022–12, 2022.

[40] Ruth Peters, Lisa Burch, James Warner, Nigel Beckett, Ruth Poulter, and Christopher Bulpitt. Haemoglobin, anaemia, dementia and cognitive decline in the elderly, a systematic review. *BMC Geriatrics*, 8:18, 2008.

[41] R C Petersen, P S Aisen, L A Beckett, M C Donohue, A C Gamst, D J Harvey, C R Jack, Jr, W J Jagust, L M Shaw, A W Toga, J Q Trojanowski, and M W Weiner. Alzheimer's

disease neuroimaging initiative (ADNI): clinical characterization. *Neurology*, 74(3):201–209, January 2010.

[42] Ann-Kristin Picke, Graeme Campbell, Nicola Napoli, Lorenz C Hofbauer, and Martina Rauner. Update on the impact of type 2 diabetes mellitus on bone metabolism and material properties. *Endocrine connections*, 8(3):R55–R70, 2019.

[43] James P Pirruccello, Alexander Bick, Minxian Wang, Mark Chaffin, Samuel Friedman, Jie Yao, Xiuqing Guo, Bharath Ambale Venkatesh, Kent D Taylor, Wendy S Post, et al. Analysis of cardiac magnetic resonance imaging in 36,000 individuals yields genetic insights into dilated cardiomyopathy. *Nature communications*, 11(1):2254, 2020.

[44] Shaun Purcell, Benjamin Neale, Kathe Todd-Brown, Lori Thomas, Manuel AR Ferreira, David Bender, Julian Maller, Pamela Sklar, Paul IW De Bakker, Mark J Daly, et al. Plink: a tool set for whole-genome association and population-based linkage analyses. *The American journal of human genetics*, 81(3):559–575, 2007.

[45] Chengxuan Qiu, Bengt Winblad, and Laura Fratiglioni. The age-dependent relation of blood pressure to cognitive function and dementia. *The Lancet Neurology*, 4(8):487–499, 2005.

[46] Yael D Reijmer, Manon Brundel, Jeroen De Bresser, L Jaap Kappelle, Alexander Leemans, Geert Jan Biessels, and Utrecht Vascular Cognitive Impairment Study Group. Microstructural white matter abnormalities and cognitive functioning in type 2 diabetes: a diffusion tensor imaging study. *Diabetes care*, 36(1):137–144, 2013.

[47] Sjur Reppe, Yunpeng Wang, Wesley K Thompson, Linda K McEvoy, Andrew J Schork, Verena Zuber, Marissa LeBlanc, Francesco Bettella, Ian G Mills, Rahul S Desikan, et al. Genetic sharing with cardiovascular disease risk factors and diabetes reveals novel bone mineral density loci. *PloS one*, 10(12):e0144531, 2015.

[48] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al. Imagenet large scale visual recognition challenge. *International journal of computer vision*, 115(3):211–252, 2015.

[49] Stephen M Smith, Gwenaëlle Douaud, Winfield Chen, Taylor Hanayik, Fidel Alfaro-Almagro, Kevin Sharp, and Lloyd T Elliott. An expanded set of genome-wide association studies of brain imaging phenotypes in uk biobank. *Nature neuroscience*, 24(5):737–745, 2021.

[50] Gary E Swan, Charles DeCarli, BL Miller, T Reed, PA Wolf, LM Jack, and D Carmelli. Association of midlife blood pressure to late-life cognitive decline and brain morphology. *Neurology*, 51(4):986–993, 1998.

[51] PGS Catalog Team. Pgs catalog calculator. `https://github.com/PGScatalog/pgsc_calc`, in preparation.

[52] Laura M Winchester, John Powell, Simon Lovestone, and Alejo J Nevado-Holgado. Red blood cell indices and anaemia as causative factors for cognitive function deficits and for alzheimer's disease. *Genome medicine*, 10(1):1–12, 2018.

[53] Ziqian Xie, Tao Zhang, Sangbae Kim, Jiaxiong Lu, Wanheng Zhang, Cheng-Hui Lin, Man-Ru Wu, Alexander Davis, Roomasa Channa, Luca Giancardo, et al. igwas: image-based genome-wide association of self-supervised deep phenotyping of human medical images. *medRxiv*, pages 2022–05, 2022.

[54] Yu Yao, Feng Gao, Yanni Wu, Xin Zhang, Jun Xu, Haiyang Du, and Xintao Wang. Mendelian randomization analysis of the causal association of bone mineral density and fracture with multiple sclerosis. *Frontiers in Neurology*, page 2049, 2022.

[55] Yau-Hua Yu, Bjorn Steffensen, Paul M Ridker, Julie E Buring, and Daniel I Chasman. Candidate loci shared among periodontal disease, diabetes and bone density. *Frontiers in Endocrinology*, 13:1016373, 2023.