Analysis_plan_GPdisc_CVD.rtf
Created: 2022-02-13 by Ida Karlsson
Updated: 2022-12-21 by Ida Karlsson
Updated: 2023-10-05 by Ida Karlsson & Elsa Ojalehto (finishing touches for archiving)

1 OBJECTIVE AND HYPOTHESES

Main objective: To study differences in the association between adiposity and CVD risk as a function of genetic predisposition to higher BMI.

When examining the association between BMI and dementia, while considering genetic predisposition to a higher BMI, we discovered that higher midlife BMI was a risk factor only among those with genetically predicted low BMI, and that higher BMI in late-life was associated with lower risk only among those with genetically predicted high BMI(1). We saw the same pattern for cognitive abilities in the Health and Retirement Study data(2), and the same has since been shown by others for mortality(3) and cardiometabolic outcomes(4), also using the Health and Retirement Study data.

This indicates that there may be differences between overweight driven by environmental factors and that driven by genetic factors. We will now continue this work, examining the effect on CVD. In addition, we will apply co-twin control models to examine additional genetic influences.

2 NOTATION AND ABBREVATIONS

CVD	Cardiovascular disease
BMI	Body mass index
GWAS	Genome-wide association study
PGS	Polygenic score
STR	Swedish Twin Registry

3 STUDY POPULATION

3.1 INCLUSION CRITERIA

Swedish Twin Registry

The STR is a population-based register of twins born in 1886-20087. In the proposed project, four sub-studies of aging will be included: GENDER and SATSA are longitudinal studies (the longest ongoing being SATSA which spanned over 30 years); HARMONY, TwinGene, and SALT-Y are cross-sectional. Only twins with genotyping data are included here, leaving study base population of 19,361 individuals (see flow chart in Appendix 1).

3.2 EXCLUSION CRITERIA

No CVD info, no adiposity info, or missing other co-variates (smoking or education). Prevalent CVD at BMI measure. Flow chart is shown in Appendix 1.

4 MEASUREMENTS AND VARIABLES

4.1 Outcome - efficacy variables

CVD

For the STR data, we use register information (National Patient Register, Cause of Death Register), with CVD defined by the ICD codes in Appendix 2. CVD is further subcategorized into non-stroke CVD and stroke.

4.2 Exposure covariates

In the STR, assessed BMI is available from each examination in the five sub-studies. In addition, height and weight at the time and at ages 25 and 40 are available from questionnaires sent out to all same-sex twin pairs in the 1960s and 70s. BMI data has been cleaned by Anna Dahl Aslan. BMI availability and cleaning is described in detail in the supplement of our 2020 BMC Medicine paper (1).

To separately examine obesity measured in midlife and late-life, we selected two different measures: That taken closest to age 50 out of measures taken between age 40 and 64 (midlife measure), and that taken closest to age 75 out of measures taken from age 65 and above.

4.3 Effect modifiers

The PGS for BMI (PGS_{BMI}) was based on GWAS summary statistics from the most recent GWAS for BMI(5), which includes around 700,000 individuals of European ancestry. The STR studies were included in the BMI GWAS, and to avoid inflation in the explanatory power of the PGS due to overlapping samples(6), new GWAS summary statistics were first generated by meta-analyzing the results, excluding those from the STR.

Genotype data were available for 19,361 individuals in the STR sample and used to compute polygenic scores (PGS). TwinGene participants (n=10,906) were genotyped on the Human OmniExpress, and SATSA, GENDER, and HARMONY (genotyped together, n=2052), and SALT-Y (n=6403) on Illumina PsychArray. The data were imputed to the HRC reference panel. The PGS computation followed the IGEMS pipeline, described on https://github.com/IGEMS/PGS_pipeline. Briefly, the PGSs were computed in Plink 1.9, after dealing with linkage disequillibrium through effect size shrinkage with SBayesR(7). The pipeline only includes HapMap3 SNPs with good imputation quality (INFO score >0.8) and MAF (>5%) on all genotyping arrays.

Prior to analyses, the PRSs were adjusted for ancestry by regressing out the first 5 PCs, and standardized within genotyping array (TwinGene; SATSA, GENDER, HARMONY; or SALT-Y), so that estimates represent the effect per standard deviation increase in PRS.

4.4 Mandatory covariates, known confounders

Age
Sex
Education (basic education ≤7 years, more than basic education >7years)
Smoking (ever versus never smoking)
Relatedness between twins (robust sandwich estimator)

4.5 Potential confounders

5 DATA MANAGEMENT

Programs, data, and related documents can be found in
P:\Dementia_IK\SfoEpi\GPdisc_CVD

6 STATISTICAL ANALYSES

All analyses were done separately for midlife and late-life measures of BMI.

We applied stratified (on study) Cox proportional hazard model with age as the timescale. Individuals were followed from the BMI measurement until first CVD diagnosis, death or end of follow-up, which was end of register follow-up for the STR sample (December 31, 2016). All models were adjusted for sex, education (basic vs more than basic education), and smoking (ever vs never smoking).

1. Independent effect model of BMI category
2. Independent effect model of the PGS_{BMI} (continuous measure)
3. Joint effect model of BMI category and the PGS_{BMI}
4. Interaction model, including an interaction between BMI category and the PGS_{BMI}
5. Stratify the results from based on tertiles of the PGS_{BMI} to test for evidence of gene-environment interactions (by including an interaction term in the model)

Using co-twin control analyses, we will:

6. Repeat model 1-5 in dizygotic twin pairs
7. Repeat model 1 in monozygotic twin pairs (model 2-5 cannot be done, as the twins will have identical PGS_{BMI})
8. Apply a modified version of model 5, by stratifying the monozygotic twin pair sample by tertiles of the PGS_{BMI}, and test model 1 in each category

Sensitivity analyses

1. Sex-stratification

Results will be presented for men and women separately as a sensitivity analysis, for both cohorts

2. Competing risk regression

After initial tests of BMI category on risk of mortalit, we will include competing risk regression as a sensitivity analysis. CVD and death will be competing events.

3. Subtypes of CVD

We will separately model stroke and non-stroke CVD, to test for subtype-specific effects (see ICD codes in Appendix 2).

7 STAFF LIST

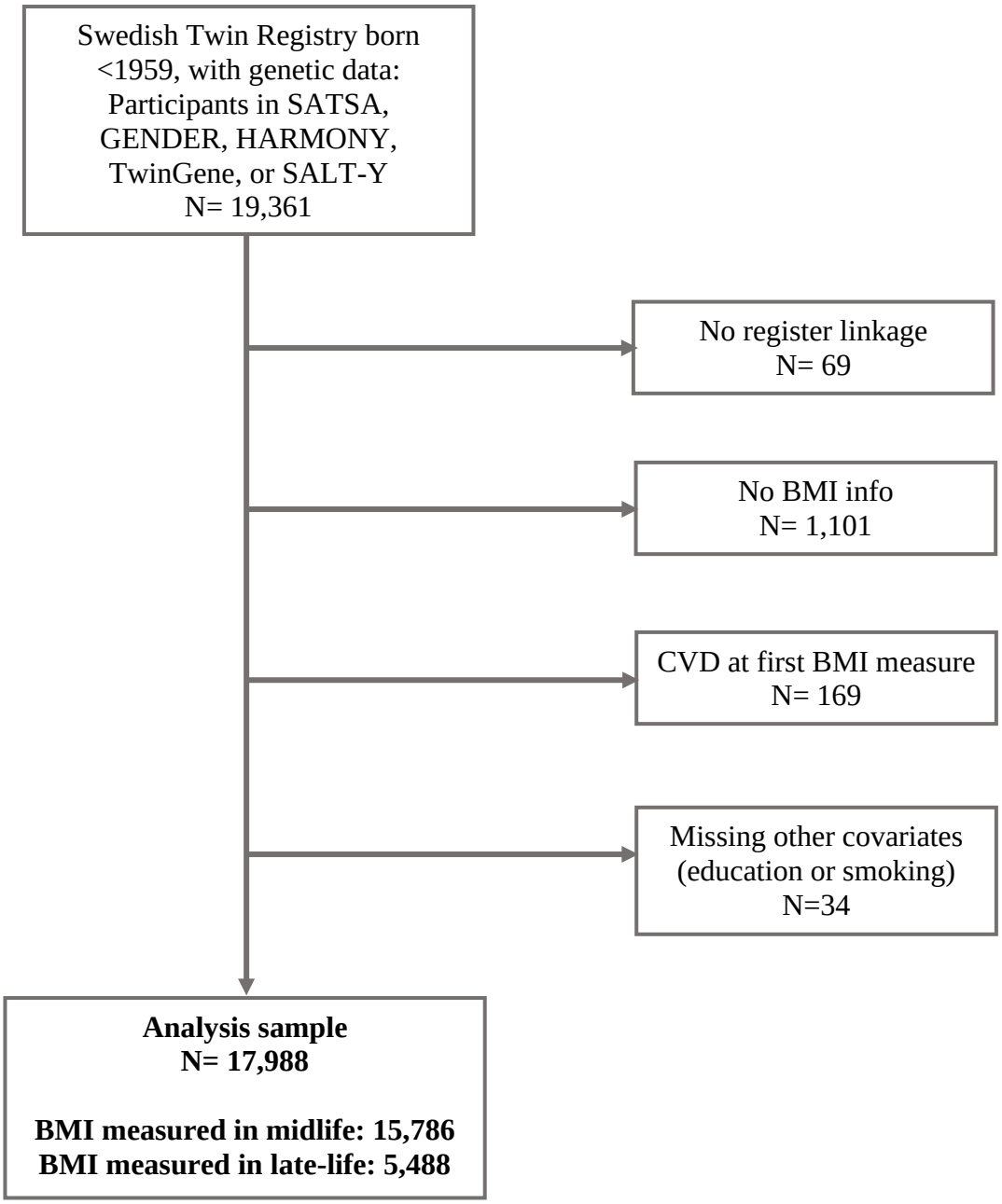
Ida Karlsson, PI
Elsa Ojalehto, first author

Yiqiang Zhan
Juulia Jylhävä
Chandra A Reynolds
Anna Dahl-Aslan
All are collaborators and co-applicants on grant applications.

References

1. Karlsson IK, Lehto K, Gatz M, Reynolds CA, Dahl Aslan AK. Age-dependent effects of body mass index across the adult life span on the risk of dementia: a cohort study with a genetic approach. BMC Medicine. 2020;18(1):131.
2. Karlsson IK, Gatz M, Arpawong TE, Dahl Aslan AK, Reynolds CA. The dynamic association between body mass index and cognition from midlife through late-life, and the effect of sex and genetic influences. Scientific Reports. 2021;11(1):7206.
3. Vinneau JM, Huibregtse BM, Laidley TM, Goode JA, Boardman JD. Mortality and Obesity Among U.S. Older Adults: The Role of Polygenic Risk. J Gerontol B Psychol Sci Soc Sci. 2021;76(2):343-7.
4. Davidson T, Vinneau-Palarino J, Goode JA, Boardman JD. Utilizing genome wide data to highlight the social behavioral pathways to health: The case of obesity and cardiovascular health among older adults. Soc Sci Med. 2021;273:113766.
5. Yengo L, Sidorenko J, Kemper KE, Zheng Z, Wood AR, Weedon MN, et al. Meta-analysis of genome-wide association studies for height and body mass index in ∼700000 individuals of European ancestry. Hum Mol Genet. 2018.
6. Wray NR, Lee SH, Mehta D, Vinkhuyzen AA, Dudbridge F, Middeldorp CM. Research review: Polygenic methods and their application to psychiatric traits. J Child Psychol Psychiatry. 2014;55(10):1068-87.
7. Lloyd-Jones LR, Zeng J, Sidorenko J, Yengo L, Moser G, Kemper KE, et al. Improved polygenic prediction by Bayesian multiple regression on summary statistics. Nat Commun. 2019;10(1):5086.
8. Ludvigsson JF, Andersson E, Ekbom A, Feychting M, Kim JL, Reuterwall C, et al. External review and validation of the Swedish national inpatient register. BMC Public Health. 2011;11:450.
9. Brooke HL, Talback M, Hornblad J, Johansson LA, Ludvigsson JF, Druid H, et al. The Swedish cause of death register. Eur J Epidemiol. 2017;32(9):765-73.

APPENDIX 1: Flow chart of the data



APPENDIX 2: CVD data from the registers (from manuscript supplement)

Healthcare registers in Sweden

The STR is connected to a number of population-based registries through the 10-digit personal identification number that is assigned to all residents in Sweden. For the current study, information on cardiovascular disease was obtained from both the National Patient Registry (NPR) and the Cause of Death Registry (CDR).

The NPR was started in 1964, and has since 1987 a coverage that contains information about 99% of all in-patient care from hospitals in Sweden (8). For each hospitalization, the primary diagnosis is recorded together with up to 20 additional diagnoses, which is done according to International Classification of Diseases (ICD) codes. The NPR also covers outpatient specialist care since 2001. In addition, all surgeries are recorded (including day surgery) according to surgical codes. The CDR has since 1961 included information about underlying and contributory causes of death for all Swedish residents, which are reported according to ICD codes as well (9). Currently, data from the NPR and CDR are available through the end of 2016. Both primary and additional diagnoses from the NPR was used together with underlying and contributing causes of death from the CDR, as criteria for disease. The following ICD codes were used to retrieve cardiovascular disease:

	ICD-7	ICD-8	ICD-9	ICD-10	Surgical code
Non-stroke CVD	420	410	410	I20	984
	450	411	411	I21	3068
	453.33	412	412	I22	3080
		413	413	I23	3127
		414	414	I24	3141
		440	440	I25	3158
		443.90	443X	I79	FNC
				I73.9	FND
					FNE
					FNG00
					FNG02
					FNG05
Stroke	330	430	430	I60	
	331.00	431	431	I61	
	331.01	433	434	I63	
	331.09	434	436	I64	
	331.99	436			
	332.00-19				
	332.29				
	334.00-98				

CVD cardiovascular disease; *ICD* International Classification of Diseases; *MI* myocardial infarction

References:

8. Ludvigsson JF, Andersson E, Ekbom A, Feychting M, Kim JL, Reuterwall C, et al. External review and validation of the Swedish national inpatient register. BMC Public Health. 2011;11:450.

9. Brooke HL, Talback M, Hornblad J, Johansson LA, Ludvigsson JF, Druid H, et al. The Swedish cause of death register. Eur J Epidemiol. 2017;32(9):765-73.