Table S5. Summary of literature review on EWAS of BMI with discovery set cohort characteristics and study design if available. All studies consist of mixed gender. None of the studies include participants with HIV. All samples came from peripheral blood except noted otherwise.

First author (Year)	(Discovery) Cohort Characteristics					Study Design				
	Size	Population	Age, Gender	BMI, Mean (SD)	Underlying conditions	Array, Sample	Study design	Validation	Statistical Analysis	Co-variables
Sharp, (2021)[58]	4894	Northern European	47.7% Female	26.98 (3.2)	NA	450K	Meta-analysis	Yes	Linear regression	Age, sex
Karlsson (2020)[32]	535	Swedish	68.2yrs; 58.5% female	26.4 (4.2)	Smoking, T2D	450K, EPIC	Cross-sectional, longitudinal	Yes	Linear mixed effects regression	GLU, CHOL, triglycerides, smoking, T2D, age, sex, methylation array
Sharma, (2020)[59]	230	African Americans	46% Female	NA	NA	450K	Cis-expression quantitative trait methylation	No	Linear regression	Age, sex, insulin sensitivity
Crocker, (2020)[60]	2325	American Indians	NA	29.6 (na)	NA	Illumina's EPIC	Population-based prospective cohort study	Yes	Linear regression	Age
He (2019)[61]	263	Penn State	16.7yrs. 44.1%female.	65.4 (28.5)	Tobacco, alcohol,	Illumina HiSeq 2500	Cross sectional	Yes	Linear regression	Age, race, sex, batch of assay
Sun (2019)[62]	1485	The Bogalusa Heart Study	44yrs; 59% female	28.7-33.3	Smoker identified	450K	Cross-sectional, longitudinal	Yes	GLM	Age, sex, smoking status, estimated WBC
Li (2019)[23]	60	CN monozygotic twins	53.53yrs; 50% female	25.1 (4.33)	NA	Illumina HiSeq X Ten	Cross-sectional, longitudinal	Yes	Linear mixed effects regression	Cell type composition, GLU, CHOL, TG, HDLC, LDLC
Campanella (2018)[63]	1941	European	NA	NA	Breast cancer, colorectal cancer, MI, b-cell malignancy, smoking	450K	Meta-analysis	Yes	Meta-analysis	Microarray and position, sex, age at blood draw, case control status
Dhana (2018)[15]	1450	Dutch	63.7yrs; 55.9%female	27.7 (4.4)	Smoking, T2D	450K	Cross-sectional	Yes	Linear mixed effects regression	Sex, age, smoking, leukocyte proportions, array number, position on array
Geurts (2018)[17]	5361	Melbourne cohort	60yrs, 32%female	NA	Prostate, colorectal, lung or kidney cancer, urothelial carcinoma or mature B neoplasms, smoking	450K	Cross-sectional, longitudinal	No	Linear mixed effects regression of case & control	Age, sex, smoking status, country of birth, sample type, white blood cell composition
Wang, (2018)[64]	700	African Americans	64% Females	NA	T2D, CVD	Sequenom's MassARRAY	Meta-analysis	Yes	Cross-sectional	Age, sex
Kvaløy, (2018)[25]	120	Norwegain	27.2yrs, 100% Female	37.04 (na)	NA	450K	Cross-sectional	Yes	Linear regression	Age
Wahl (2017)[1]	5387	European, Asian Indian	>50yrs. >50%female	>26.8	Alcohol, hypertension, coronary heart disease, T2D, smoking	450K	Cross-sectional, longitudinal	Yes	Inverse variance meta-analysis	Age, sex, smoking, physical activity, alcohol, probe, estimated WBC proportions

Study	N	Cohort/Ancestry	Age/Sex	BMI	Conditions	Array/Tissue	Design		Method	Adjustments
Sayols-Baixeras (2017)[4]	641	REGICOR, FOS	63.2yrs, 50.7% female	27.0 (4.0)	Hypertension, diabetes, smoking,	450K	Cross-sectional	Yes	Fixed effects meta-analysis of two cohorts	Age, sex, smoking, surrogate var
Mendelson (2017)[65]	3743	FHS & LBC	67yrs,55% female	28.3 (5.4)	Diabetes, coronary artery disease.	450K	Cross-sectional	Yes	Meta-analysis of two cohorts	Age, sex
Meeks (2017)[8]	547	African	50.5yrs, 57.8%female	26.7 (0.5)	Without T2D	450K	Cross-sectional	Yes	Linear regression	Age, sex, recruitment site, estimated cell distributions, technical effects, first principal component from genotyping data
Crujeiras (2017)[66]	55	European Caucasian	27.4yrs; 23% female;	NA	Diabetes	450K; adipose tissue, leukocytes	Case control	No	Wilcoxon rank test	Age
Ali (2016)[67]	192	Northern European ancestry	36.2yrs, 55%females	NA	Insulin resistance, hypertriglyceridemia	450K	bisulfite validation seq	Yes	SOLAR, rank-normal transformation	Sex, age, interactions
Al Muftah, (2016)[68]	123	Middle Eastern	39yrs, 58%female	28.3 (6.2)	T2D	450K	Meta-analysis	Yes	Linear mixed model	T2D, BMI, age, sex
Demerath (2015)[16]	2097	African	56.2yrs, 64%female	30.1 (6.1)	Smoking, alcohol	450K	Cross-sectional	Yes	Linear mixed effects regression	Leukocyte proportions, sex, age, study center, WBC, education, household income, smoking, alcohol, physical activity, array
Aslibekyan (2015)[6]	991	GOLDN	49yrs; 52%female	28 (6)	Smoking	450K	Cross sectional; bisulfite sequencing	Yes	Linear mixed effects regression	Age, sex, smoking status, T cell purity, study site, family structure
Wang, (2015)[69]	220	Chinese Han	NA	28.4 (3.8)	NA	Sequenom's MassARRAY	Case-control study	Yes	Liner regression	Age
Dick (2014)[10]	479	European	43.8yrs, 78%female	24.2 (na)	MI, smoking, diabetes,	450K; skin, adipose tissue	Cross-sectional	Yes	Linear mixed effects regression	age, sex, smoking, methylation array batch, MI