The genetic case for cardiorespiratory fitness as a clinical vital sign and the routine prescription of physical activity in healthcare

Ken B. Hanscombe[1, 4], Elodie Persyn[1], Matthew Traylor[2], Kylie P. Glanville[4], Mark Hamer[3], Jonathan R. I. Coleman[4], Cathryn M. Lewis[1, 4]

[1]Department of Medical & Molecular Genetics, King's College London, London, United Kingdom, [2]Queen Mary University of London, London, United Kingdom, [3]Institute of Sport Exercise & Health, Division of Surgery and Interventional Science, University College London, London, United Kingdom, [4]Social, Genetic and Developmental Psychiatry Centre, King's College London, London, United Kingdom

Ken B. Hanscombe ken.b.hanscombe@kcl.ac.uk (corresponding author)
Elodie Persyn elodie.persyn@kcl.ac.uk
Matthew Traylor m.traylor@qmul.ac.uk
Kylie P. Glanville kylie.glanville@kcl.ac.uk
Mark Hamer m.hamer@ucl.ac.uk
Jonathan R. I. Coleman jonathan.coleman@kcl.ac.uk
Cathryn M. Lewis cathryn.lewis@kcl.ac.uk

Abstract

Background: Cardiorespiratory fitness (CRF) and physical activity (PA) are well-established predictors of morbidity and all-cause mortality. However, CRF is not routinely measured and PA not routinely prescribed as part of standard healthcare. The American Heart Association (AHA) recently presented a scientific case for the inclusion of CRF as a clinical vital sign based on epidemiological and clinical observation. Here, we leverage genetic data in the UK Biobank (UKB) to strengthen the case for CRF as a vital sign, and make a case for the prescription of PA.

Methods: We derived two CRF measures from the heart rate data collected during a submaximal cycle ramp test: CRF-vo2max, an estimate of the participants' maximum volume of oxygen uptake, per kilogram of body weight, per minute; and CRF-slope, an estimate of the rate of increase of heart rate during exercise. Average PA over a 7-day period was derived from a wrist-worn activity tracker. After quality control, 70,783 participants had data on the two derived CRF measures, and 89,683 had PA data. We performed genome-wide association study (GWAS) analyses by sex, and post-GWAS techniques to understand genetic architecture of the traits and prioritize functional genes for follow-up.

Results: We found strong evidence that genetic variants associated with CRF and PA influenced genetic expression in a relatively small set of genes in heart, artery, lung, skeletal muscle, and adipose tissue.

These functionally relevant genes were enriched among genes known to be associated with coronary artery disease (CAD), type 2 diabetes (T2D), and Alzheimer's disease (three of the top 10 causes of death in high-income countries) as well as Parkinson's disease, pulmonary fibrosis, and blood pressure, heart rate, and respiratory phenotypes. Genetic variation associated with lower CRF and PA was also correlated with several disease risk factors (including greater body mass index, body fat and multiple obesity phenotypes); a typical T2D profile (including higher insulin resistance, higher fasting glucose, impaired beta-cell function, hyperglycaemia, hypertriglyceridemia); increased risk for CAD and T2D; and a shorter lifespan.

Conclusions: Genetics supports three decades of evidence for the inclusion of CRF as a clinical vital sign. Given the genetic, clinical, and epidemiological evidence linking CRF and PA to increased morbidity and mortality, regular measurement of CRF as a marker of health and routine prescription of PA could be a prudent strategy to support public health.

Keywords

cardiorespiratory fitness, physical activity, genetics, genome-wide association, UK Biobank

Background

Cardiorespiratory fitness (CRF) and physical activity (PA) are well-established predictors of morbidity and mortality (1, 2). CRF is such a strong predictor of cardiovascular health, all-cause mortality, and mortality attributable to various cancers, that in 2016 the American Heart Association (AHA) put forward a scientific case for measuring CRF as a clinical vital sign (3). The AHA recommended that, "(at) a minimum, all adults should have CRF estimated each year (…) during their annual healthcare examination". Non-exercise testing methods for estimating CRF make this possible in a clinical setting (4). The AHA also noted that measurement of CRF provides "clinicians the opportunity to counsel patients regarding the importance of performing regular physical activity". That exercise is medicine, preventative and curative, has been known from antiquity – Hippocrates wrote the earliest known prescription for exercise about 2000 years ago (5). Today however, we do not yet routinely prescribe PA, nor do we measure CRF as part of a standard clinical assessment.

Sedentary lifestyles are a significant public health problem. For example, physical inactivity and poor diet were the second leading modifiable, behavioural cause of death in the US in 2000 and are expected to overtake tobacco to become the leading cause (6). Worldwide, physical inactivity was estimated to account for about 10% of premature mortality in 2008, an effect size similar to smoking and obesity (7). In 2013, the economic burden due to direct health-care costs, productivity losses, and disability-adjusted life-years attributable to physical inactivity was estimated at US$53.8 billion (8). This economic burden is reflected in the prevalence of physical inactivity worldwide: 28% in 2016 (23% for men, 32% for women) (9). Most people

do not know what the recommended activity guidelines to maintain good health are. A UK British Heart Foundation (BHF) survey estimated about 60% of adults in the UK do not know the recommended minimum level of PA (10). Modifiable physiological characteristics like CRF and behaviours like regular exercise are affected by both environmental and genetic factors. Genetic factors, not addressed in the AHA review, can influence CRF and PA through multiple mechanisms including appetite for leisure activity, physiological response to exercise, and CRF response to PA. Critically, genetics can be leveraged to identify links between CRF, PA, and disease. Here we present the genetic case for the measurement of CRF as a clinical vital sign, and the routine prescription of PA.

CRF is typically measured with incremental exercise (resistance or workload) on a treadmill or cycle ergometer. In a sub-maximal test, CRF is estimated from extrapolating the workload to the age-estimated maximum heart rate. PA has most often been measured by self-report. Previous research has shown that CRF and PA are familial, with family and twin study estimates of their heritability varying widely (CRF: h^2=0.25 – 0.65 (11); PA: h^2=0 – 0.78, and 0.48 in the only study that objectively measured PA (12)). There has been some interest in variation in CRF response to exercise. Several candidate gene studies and one genome-wide association study (GWAS) have reported genetic variants associated with baseline CRF, or CRF change in response to exercise, but none was genome-wide significant or reliably replicated (13-15). Two GWASs of PA in the UK Biobank (UKB) reported three loci associated with accelerometer-measured 7-day average PA (SNP heritability, h^2_{SNP}=0.14), and none with self-reported PA (16, 17). The genetic contribution to CRF is poorly characterised, and our current understanding of the genetics of PA has not examined differences by sex.

In this study, we performed the largest genome-wide association analysis of CRF, a GWAS of PA, and related their genetic risk profiles to the known genetic risk for anthropomorphic phenotypes, metabolic traits, and chronic diseases. CRF and PA were objectively measured in the UKB (18). Given the multitude of mechanisms by which PA and CRF affect biology (3, 15, 19), the differences in physiognomy (e.g., body fat distribution (20, 21)), behavioural dynamics (e.g., motivators and context preferences (22)), we investigated sex-specific as well as sex-combined CRF and PA phenotypes. Previous UKB analyses of PA have reported genetic correlations between the sexes and included sex as a covariate in combined analyses (17), which by design identifies only genetic factors common to both sexes. Here, we also investigated the potential functional effects of sex-specific associations. Our aim in this study was to characterise the genetic contribution to CRF and PA highlighting genetic links to disease, and thus to strengthen the case for the inclusion of CRF as a clinical vital sign and make more cogent the argument for routine prescription of regular PA in healthcare.

Methods

Sample

UK Biobank. The UK Biobank (UKB) is a research study of over 500,000 individuals from across the United Kingdom. Participants aged 40 to 69 were invited to attend one of 22 assessment centres between 2006 and 2010. A comprehensive range of health-related data has been collected at baseline and follow up, including disease diagnoses and lifestyle information, with blood samples taken for genome-wide genotyping and a biomarker panel. Further data are available in follow up studies on different subsets of UKB participants, including CRF and PA data. All participants provided electronic signed consent (18, 23).

After selecting the European subsample (the largest cluster from a 4-means clustering of the first 2 UKB-derived principal components), and phenotype and genotype QC described below, we had 70,783 (34,419 male, 36,364 female) participants with CRF data, and 89,683 (39,352 male, 50,331 female) participants with PA data. Additional file 1: **Table S1** shows details of the derivation of the final GWAS sample.

Phenotypes

Cardiorespiratory fitness - submaximal cycle ramp test.

CRF was assessed with a submaximal cycle ramp test which used a stationary bicycle and 4-lead electrocardiograph (ECG) to record heartrate data during pre-test, exercise and rest phases (approximately 15 seconds, 6 minutes, and 1 minute respectively). The cycle test procedure is described in detail in the UKB Cardio Assessment document (24). Cycle ramp test data were collected at two instances: from August 2009 at the end of the initial assessment visit 2006–2010 (instance 0), and the first repeat assessment visit 2012–2013 (instance 1). Participants performed an exercise protocol determined by their risk level, *Program Category* (UKB field 6024): 'minimal', 'small', 'medium', or 'high'. The cycle ramp test included a 'Pre-test', 'Exercise', and 'Recovery' phase. We used data from the 'Exercise' phase only, within which, only 'minimal' and 'small' risk categories included a workload ramp (increase in workload). We selected samples in the 'minimal', 'small', and 'medium' risk categories, with *ECG Bike Method for Fitness Test* (6019) = 'Bicycle' (Additional file 1: **Table S2**), modelling heart rate data from the 4-minute workload ramp for the 'small' and 'minimal' risk groups, and the full 6-minute constant work load for the 'medium' risk group. We also included risk category as a covariate in the genetic association analyses.

The raw heart rate data were noisy (Additional file 1: **Figure S1**, panel **A**, shows raw heart rate data for 1,000 random participants). Participants were instructed to cycle at a cadence of 60 revolutions per minute (RPM) and we selected workload and heart rate data where participants cycled in the range 35-125 RPM (Additional file 1: **Figure S1**, panel **B**), within the manufacturer guaranteed range for cycle ergometer workload accuracy. Finally, we selected samples with a minimum of 20 observations of heart rate, workload, and cadence within the 'Exercise' phase of the test, and applied a Butterworth filter to the heart rate data (**Figure 1**). Full cycle ramp test QC is described in the supplementary Methods and listed in Additional file 1: **Table S1**.

From this data, we derived two measures: *CRF-vo2max*, an estimate of the maximum volume of oxygen uptake per kilogram of body weight per minute; and *CRF-slope*, the rate of increase of heart rate. Specifically, we fit a linear model to workload and heart rate and extrapolated from the linear model to predict workload at the age-estimated maximum heart rate (HR·age·max = 208 – 0.7 * age, where age is age when attended assessment centre). Maximum workload relative to body weight in kilograms, or relative power (watts/kg), was then used as a proxy for CRF-vo2max:

$$\text{CRF-vo2max} = \text{WL·HR·max} \ / \ \text{weight}$$

Where WL·HR·max is extrapolated work load at age-estimated maximum heart rate. The coefficients 208 and 0.7 for age-estimated maximum heart rate are commonly used constants and were used in previous studies of CRF in the UKB (25-28).

For CRF-slope, we fitted a linear model to the Butterworth-filtered heart rate signal for each participant and retained the slope, β_1, as rate of increase of heart rate (**Figure 1**):

$$\text{HR} \sim \beta_0 + \beta_1 * \text{trend}$$

Where HR is heart rate and 'trend' refers to the time points at which repeated measurements of heart rate, workload, and cadence were taken within each phase of the submaximal cycle ramp test (**Figure 1**). Samples with negative slopes were excluded.

Physical activity - wrist-worn accelerometer.

PA in the UKB was measured continuously over a period of 7 days with a wrist-worn accelerometer in 103,712 participants, as an add-on measure after the initial assessment between June 2013 and January 2016. We used a wear-time adjusted 7-day average PA and filtered participants on the following test-specific variables: *Data quality, good wear-time* (90015)='Yes', *Data quality, good calibration* (90016)='Yes', and no *Data problem indicator*s (90002). Finally, we excluded samples with PA (accelerometer) > 100 milligravities (mg). General PA quality control and calibration of raw accelerometer data and calculation of derived activity level are described in detail elsewhere (29).

Disease frequency

The top 10 global causes of death in high-income countries in 2016, as reported by the World Health Organization (WHO), are shown in Additional file 1: **Table S3**, along with the relevant International Classification of Diseases revision 10 (ICD-10) codes. Ischaemic heart disease was the number 1 global cause of death among high-income countries; lower respiratory tract infections, ranked 6th, was the only communicable cause of death on the list. We used UKB in-patient hospital episode statistics (diagnoses recorded in ICD-10 codes) to retrieve disease frequency for the top 10 global causes of death, among the

subsample with CRF or PA data (Additional file 1: **Table S3** includes ICD-10 codes). We ranked participants by level of the phenotypes CRF-vo2max, CRF-slope, and PA after correcting for age and socioeconomic status (SES, Townsend deprivation index), divided the participants into ten approximately equal-sized groups, and calculated the proportion in each group with a main or secondary diagnosis (made during a hospital in-patient admission) for each of the 10 query diseases. ICD-9 diagnoses were not included. Firstly, it is not clear how to map ICD-9 codes onto the ICD-10 codes provided by the WHO. Secondly, ICD-9 codes were observed at very low frequencies: 1 sample with CRF data, and 40 with PA data, were diagnosed with ischaemic heart disease (ICD-9 codes starting with 41[0-9]), the most prevalent of the top 10 global causes of death. Diagnoses were potentially received both before and after the cycle ramp test and the week in which participants wore the physical activity tracker, and we made no distinction between incident and prevalent disease. This analysis was performed separately in males and females, since there are differences in both disease prevalence, and in phenotype distribution by sex.

Genetic analyses

The full UKB sample were genotyped on one of two custom Applied Biosystems genotyping arrays: 49,950 on the UK Biobank Lung Exome Variant Evaluation (UK BiLEVE) Array, and 438,427 on the UK Biobank Axiom Array. The genetic data underwent standard GWAS QC and were imputed using IMPUTE4 to three reference panels by the UKB: the Haplotype Reference Consortium (HRC) and the merged UK10K and 1000 Genomes phase 3 reference panels. General genotyping considerations, raw genotype data QC, and genetic imputation procedures in the UKB are described in detail elsewhere (18). Details of available genetic data and associated metadata are described in UKB Resource 664 (30).

From UKB-supplied sample QC information, we excluded gender mismatches, samples with putative sex chromosome aneuploidy or excess relatives, and heterozygosity and missingness outliers. We performed linear mixed model association on the residual of each trait regressed on the covariates age, age^2, sex (for sex-combined analyses), array, centre, and the first 10 principal components calculated on subsamples with data on the trait of interest. Additionally, for CRF phenotypes, we included the covariates risk category (corresponding to CRF test protocol) and number of trend entries (number of observations). Full details of GWAS QC and association parameters, alternative models including BMI, and alcohol and smoking as covariates, and post-GWAS analyses including transcriptome-wide association (TWAS) and colocalization, gene-set enrichment, and genetic correlation, are included in the supplementary methods. We used the standard $p=5x10^{-8}$ as our genome-wide significant threshold, a conservative Bonferroni-corrected p-value for TWAS significance correcting for all genes tested in all tissues ($p=0.05/75,486=6.6x10^{-7}$), and a false-discovery rate (FDR) correction for gene-set enrichment and genetic correlation accepting FDR-adjusted p-values<0.05 as significant.

Analysis tools

For data analysis and figure creation we used both R (31) and Python 3 (32), in conjunction with the bioinformatic pipeline software Snakemake (33). We used the R package ukbtools (34) to parse all UKB data. Genetic data analyses were performed using FlashPCA (35), BOLT-LMM (36), PLINK 1.9 and 2.0 (37) , LDSC (38), FUSION (39), and the web-based tools FUMA (40), LD Hub (41), and LDlink (42).

Results

Descriptive statistics and smoking and alcohol drinking status

Descriptive statistics for the CRF and PA phenotypes showed higher values for CRF-vo2max in males, and higher CRF-slope and PA in females (Additional file 1: **Table S4**). Correlations between CRF and PA phenotypes, and age and BMI were all modest and negative (r=-0.02 – -0.27), except for CRF-slope and age which had a small positive correlation in females (r=0.13, p=1.2x10^{-144}) (Additional file 1: **Figure S2**). Combined-sex sample correlations between CRF-vo2max and CRF-slope, and SES were not significant; the PA and SES, and sex-specific correlations for all three traits were significant but small ($|r|$<=0.03). In the combined-sex sample, phenotypic correlations between our three traits were in the directions expected. CRF-vo2max and PA had a modest positive correlation (r=0.25, p=1.5x10^{-195}, among the subsample who performed both the submaximal cycle ramp test and participated in the 7-day PA monitoring, N=13,561). CRF-slope and PA were not significantly correlated (r=-0.01, p=0.11, N=13,561). The two fitness phenotypes, CRF-vo2max and CRF-slope were negatively correlated (r=-0.39, p<0.001, N=70,783).

As a proxy for true resting heart rate, heart rate at the beginning of the exercise phase of the cycle ramp test was negatively correlated with CRF-vo2max (r=-0.33, p<0.001), and had a near zero correlation with CRF-slope (r=-0.03, p=3.8x10^{-13}).

Smoking and alcohol drinking status are sometimes considered confounding factors in analyses of CRF and PA. In our data, there were small differences in mean CRF-vo2max, CRF-slope, and PA by smoking and drinking status (Additional file 1: **Table S5**), which contributed little to trait variance in linear regression models of all three phenotypes (ΔR^2<=0.005 compared to a base model of age and sex, Additional file 1: **Table S6**). We chose not to adjust for the effects of smoking and drinking status in order to obtain unbiased estimates of the genetic factors on CRF and PA. There is evidence that correcting for heritable covariates can lead to biased estimates of the influence of genetic factors (43).

Disease frequency

For the top 10 global causes of death in high-income countries in 2016, as reported by the World Health Organization (WHO), **Figure 2** shows the frequency at which each disease was observed in the UKB by level of CRF-vo2max, CRF-slope, and PA (after correcting for age and SES). Given the difference in the

phenotype distributions of the male and female subsamples, and the inherent difference in disease frequency by sex, disease frequency has to be interpreted separately by sex.

Disease frequency by CRF

With increasing levels of CRF-vo2max, there were reductions in diagnosis frequency for ischaemic heart disease (CAD) (male: 12.0–6.1%, female: 4.9–1.5%), stroke (male: 1.5–0.7%, female: 0.6–0.4%), chronic obstructive pulmonary (COPD) (male: 1.5–0.3%, female: 1.1–0.3%), lower respiratory tract infection (LRTI) (male: 2.5–1.0%, female: 2.0–1.2%), diabetes mellitus (male: 5.0–1.8%, female: 2.4–0.7%), and kidney diseases (male: 1.4–0.8%, female: 1.0–0.6%). The largest differences were observed for CAD and diabetes, with the starkest reductions for all disease often seen within the lowest 4 deciles of CRF-vo2max.

CRF-slope showed reduction in disease for CAD (male: 9.2–5.6%, female: 2.6–1.5%) and diabetes (male: 4.5–2.4%, female: 1.5–0.6%). The relationships with other disease diagnoses were more noisy.

Disease frequency by PA

Disease diagnosis by level of PA showed similar patterns to CRF-vo2max, with a continued decile-to-decile reduction in disease across the range of PA from least to most physically active. With increasing PA we observed reductions in CAD (male: 9.3–2.5% female: 3.0–1.0%), stroke (male: 2.1–0.4%, female: 1.0–0.2%), COPD (male: 1.8–0.2%, female: 1.1–0.3%), LRTI (male: 2.8–0.8%, female: 2.0–1.0%), diabetes mellitus (male: 6.0–0.7%, female: 3.7–0.4%), and kidney diseases (male: 2.4–0.9%, female: 1.3–0.5%). Like CRF–vo2max, the greatest reductions were seen in CAD and diabetes mellitus.

Genetic associations

Genome-wide association studies were performed for CRF-vo2max, CRF-slope, and PA with results for male, female, and the combined sample (**Figure 3**). An annotated list of the independent genome-wide significant variants, determined by LD clumping of the association results, is included in **Table 1**. Additional file 1: **Figures S3-S37** are the regional plots for each sex-combined GWAS signal peak, and sex-specific GWAS signal peaks where there was evidence of heterogeneity.

<insert **Table 1** here>

Sex-combined results

We found twelve significant SNPs for CRF-vo2max and eight independent SNPs for CRF-slope, in the combined sample (**Table 1**). For PA, we found nine significant SNPs. However, three of these were on chromosome 17, two of which are situated within the large 17q21.31 inversion polymorphism (43,624,578 – 44,525,051 MB, human genome build 37 coordinates) and may not reflect independent causal variants (44,

45). We replicated two of the three previously reported PA GWAS significant SNPs (17), rs2696625 (chr17) and rs59499656 (chr18) , and tagged the third rs564819152 (chr10) through LD with rs34719019 (chr10), r^2=0.78 in 1000 Genomes EUR (Additional file 1: **Table S7**).

Sex-specific results

For CRF-vo2max we found four male genome-wide significant SNPs, three of which showed no evidence of heterogeneity and were associated (directly or indirectly through LD) in the combined sample: rs9809798 (r^2=0.79 with rs6801957, chr3), rs111299422 (chr5), rs1006545 (r^2=0.97 with rs11190709, chr10). rs41317306 (chr14) however, showed no evidence for association in females (p=5.1x10^{-1}), or the combined sample (p=1.6x10^{-4}), and significant effect size heterogeneity (Q=22.13, p<0.001). However, the regional plot for rs41317306 did not show support from correlated SNPs (Additional file 1: **Figure S11**)

For CRF-slope, the one female signal rs111299422 (chr5) was significant in the combined sample, but showed evidence of heterogeneity (Q=8.37, p=3.8x10^{-3}). There were two male specific SNPs: rs1006545 (chr10) was indirectly associated in the combined sample (r^2=0.97 with 10:102554618_AT_A) and showed no evidence of heterogeneity. rs1741294 (chr20) showed no evidence of association in females (p=6.0x10^{-1}) or the combined sample (p=4.2x10^{-4}), and significant effect size heterogeneity (Q=7.96, p=4.8x10^{-3}).

For PA, there were four SNPs associated in the female sample, two of which were associated in the combined sample, rs62055696 (chr17) and 17:44262581_A_C (r^2=0.70 with rs62055696, chr17), and showed no evidence of heterogeneity. The two female signals on chromosome 16 signals rs75986475 and rs13329850 were uncorrelated (r^2=0.0005), showed significant heterogeneity (Q=8.51, p=3.5x10^{-3}; Q=16.98, p<0.001 respectively), and no evidence of association in males (p=3.3x10^{-1}, p=4.5x10^{-1} respectively) or the combined sample (p=1.3x10^{-6}, p=7.6x10^{-4} respectively). The male-specific SNP identified in PA, rs78661713 (chr2), showed no evidence for association in females (p=6.1x10^{-1}) or the combined sample (p=5.9x10^{-4}), and significant heterogeneity (Q=20.63, p<0.001).

Genes to prioritize for experimental follow-up

For each trait, we report genes identified as both *functional* candidates by TWAS and colocalization analysis, and *positional* candidates (**Table 1**) or gene-based association candidates (Additional file 1: **Table S8**). We then report any additional noteworthy TWAS-significant functional candidates. The functional candidates reported below all passed the TWAS multiple testing threshold (16 for CRF-vo2max, 10 for CRF-slope, and 24 for PA), and are shown in **Figure 4**. Where functional candidates were also implicated in the gene-based association test these are indicated in **Figure 4**. The full list of significant genes from the gene-based test of association are shown in Additional file 1: **Table S8**.

CRF-vo2max

All functional candidates were located within cytochrome band 17q21.31, within or near the chromosome 17 inversion. In particular, five functional candidates were identified in the gene-based association and had a high probability of a shared SNP driving both the CRF-vo2max association and expression in the listed tissues (PP4 range=0.94 – 0.99, **Figure 4**): *WNT3* (Wnt family member 3) in aorta, heart (atrial appendage), lung and skeletal muscle; *MAPT* (microtubule associated protein tau) in brain and lung; *LRRC37A2* (leucine rich repeat containing 37A2) in and *LRRC37A* (leucine rich repeat containing 37A) in adipose (subcutaneous and visceral omentum), artery (aorta and tibial), heart (atrial appendage and left ventricle), and skeletal muscle; and *CRHR1* (corticotropin releasing hormone receptor 1) in adipose (subcutaneous and visceral omentum) and skeletal muscle. Notable GWAS catalog associations for this set of genes include disease (atrial fibrillation (46), diabetes mellitus, and coronary artery disease (47), Parkinson's (48), and Alzheimer's disease (49)), lung function (FEV1, FVC (50), peak expiratory flow (51)), heart and electrocardiography measures (QRS measures (52), cardiac arrhythmia (53)), and blood phenotypes (systolic blood pressure, eosinophil count (54), haemoglobin level (55), haematocrit (56), red blood cell density (57)).

CRF-slope

Positional candidate *ACHE* (acetylcholinesterase), implicated by rs4582488 (52.87kb upstream), was significantly associated with expression in adipose (subcutaneous, visceral omentum), artery (aorta, coronary, and tibial), and heart (atrial appendage) tissue. Each of these tissue associations had high probability of a shared SNP driving both the CRF-slope association and expression (PP4 range=0.79 – 0.84). *ACHE* has notable reported associations with resting heart rate (58), heart rate response to exercise and recovery (59), electrocardiography (60), diastolic blood pressure (61), and diabetes mellitus (62).

Three functional candidates were also significant in the gene-based association: *TRIP6* (thyroid hormone receptor interactor 6) with high probability of a shared SNP driving both CRF-slope association and expression in whole blood, and *PINX1* (PIN2 (TERF1) interacting telomerase inhibitor 1) and *ERI1* (exoribonuclease 1) had significant expression in lung, with strong evidence the same variant drives expression and the CRF-slope association (PP4 range=0.71 – 0.98). *TRIP6*, *PINX1*, and *ERI1* have notably been associated with blood pressure (63), triglyceride levels (64), HDL cholesterol (65), diabetes mellitus (66), and atherosclerosis (67).

Additional TWAS significant functional candidates with high probability of the GWAS variant driving expression included *SRRT* (serrate, RNA receptor molecule) in lung and heart (atrial appendage) tissue, *RP11-62H7.2* in heart (atrial appendage), *FAM86B3P* (family with sequence similarity 86 member B3) in skeletal muscle, *FAM85B* (family with sequence similarity 85 member B) in lung and coronary artery, and *AF131215.2* in artery (coronary and tibial) (PP4 range=0.74 – 0.92). *SRRT* and *FAM86B3P* are notably associated with heart rate response to exercise and recovery (59), electrocardiography (RR interval) (68), diabetes mellitus (66), and obesity (69).

PA

All PA functional candidates were on chromosome 17, except *L3MBTL2* on chromosome 22. The number of loci and the best candidates should be interpreted with caution as this is a gene dense region with many copy number variants, and all the chromosome 17 candidates were within or near the large 17q21.3 inversion polymorphism (44, 45). Nonetheless, the positional candidates were supported by the functional analyses.

Ten significant genes from the gene-based association analysis were identified as functional candidates: *PLEKHM1* (plekstrin homology domain containing M1) in lung, skeletal muscle, and peripheral blood with low probability of the GWAS variant driving expression (PP4 range=0.30 – 0.33); *NSF* (N-Ethylmaleimide-Sensitive Factor) in coronary artery (PP4=0.54); *MAPT* in subcutaneous adipose (PP4=0.91), artery (aorta PP4=0.96, tibial PP4=0.67), and brain, left ventricle, lung, adipose (PP4 range=0.93 – 0.99); *LRRC37A2* in adipose (visceral omentum and subcutaneous), artery (aorta, coronary), heart (atrial appendage and left ventricle), lung (all PP4=0.99); *KANSL1* (KAT8 Regulatory NSL Complex Subunit 1) in skeletal muscle (PP4=0.92); *CRHR1* in adipose (visceral omentum and subcutaneous) and skeletal muscle (PP4>=0.98); *ARL17A* (ADP ribosylation factor like GTPase 17A) with high probability of a shared variant driving PA association and gene expression in left ventricle (PP4=0.88) and skeletal muscle (PP4=0.97); *ARHGAP27* (rho GTPase activating protein 27) in visceral omentum (PP4=0.92); and *L3MBTL2* (L3MBTL histone methyl-lysine binding protein 2) in subcutaneous adipose (PP4=0.67).

Previous notable associations with these ten functional candidates include disease (Parkinson's disease (70), hypertrophic cardio myopathy (71), ischaemic stroke, diabetes mellitus and coronary artery disease (72)) anthropometrics (BMI-adjusted waist hip to ratio, BMI (73)), blood phenotypes (haemoglobin measurement (55), red blood cell density, hematocrit (57), eosinophil count and systolic blood pressure (54)), lung function (FVC (54), FEV1 (51)), and heart and electrocardiography measures (QRS measures (52), cardiac arrhythmia (53)).

Figure 4 includes the full list of TWAS significant candidates, including those with lower posterior probabilities of a shared variant driving both GWAS trait association and expression.

Gene set enrichment

We also performed gene set enrichment analyses on the TWAS significant genes for sex-combined CRF-vo2max (16 genes), CRF-slope (10 genes) and PA (24 genes, 23 with recognized Ensembl ID), as well as female PA (22 genes: 21 with recognized Ensembl ID). TWAS analyses of the other sex-specific traits did not yield sufficient gene associations to perform enrichment analyses. Enrichment of TWAS-identified functional candidates in gene sets defined by genes identified in previous GWASs, reported in the GWAS Catalog, are shown in Additional file 1: **Figures S38**–**S40**.

Functional candidates for CRF-vo2max were enriched in GWAS gene sets for multiple diseases including multiple system atrophy, Alzheimer's disease, Parkinson's disease, idiopathic pulmonary fibrosis, and primary biliary cirrhosis (FDR-adjusted p range=$6.9x10^{-3}$ – $2.0x10^{-5}$), as well as lung function (FEV1, FDR-adjusted p=) and blood phenotypes (haematocrit FDR-adjusted p values=$2.8x10^{-2}$, haemoglobin concentration FDR-adjusted p values= $4.0x10^{-2}$). Enrichment for genes implicated in Alzheimer's and Parkinson's disease are noteworthy given there is some evidence of an underlying insulin resistance pathology.

CRF-slope functional genes were enriched in GWAS gene sets for the thrombosis and atherosclerosis risk factor *PAI-1* (plasminogen activator inhibitor-1: FDR-adjusted p=$5.0x10^{-6}$), heart rate response to exercise and recovery (FDR-adjusted p=$1.2x10^{-5}$), and resting heart rate (FDR-adjusted p=$1.4x10^{-3}$)

GWASs genes sets enriched for PA functional candidates included Parkinson's disease, idiopathic pulmonary fibrosis, Alzheimer's disease, multiple system atrophy and primary biliary cirrhosis (FDR-adjusted p range=$2.7x10^{-2}$ – $2.0x10^{-8}$), and mental health traits (e.g., neuroticism, mood instability, and alcohol use disorder). Also noteworthy were enrichment of female PA-associated functional candidates in FEV1 (FDR-adjusted p=$4.1x10^{-4}$) and FVC (FDR-adjusted p=$3.1x10^{-3}$) given their use in the diagnosis of obstructive and restrictive pulmonary diseases.

Genetic correlations within CRF-vo2max, CRF-slope, and PA

CRF-vo2max, CRF-slope and PA were all polygenic with modest SNP heritabilities of between 0.08 and 0.14 (**Table 2**). There was no observed difference between sexes.

Table 2. Heritability of CRF and PA

	CRF-vo2max	CRF-slope	PA
Combined	0.10 (0.01)	0.08 (0.01)	0.14 (0.01)
Male	0.10 (0.01)	0.09 (0.02)	0.13 (0.01)
Female	0.11 (0.02)	0.10 (0.02)	0.15 (0.01)

Note. Values shown are SNP heritabilities and standard errors on the observed scale. CRF-vo2max = maximum volume of oxygen uptake per kg of bodyweight per minute, estimated from workload and heart rate during the cardiorespiratory fitness test, as well as age and weight; CRF-slope = the rate of increase of heart rate during the exercise period of the cycle ramp test; PA = 7-day average physical activity

For the sex-combined between-phenotype genetic correlations, CRF-vo2max and CRF-slope were negatively correlated (r_g=-0.66, se=0.04, p=$5.4x10^{-53}$), suggesting genetic variants that contribute to greater fitness are also associated with a lower rate of increase of heart rate during exercise. PA had a moderate positive correlation with CRF-vo2max (r_g=0.37, se=0.04, p=$1.9x10^{-16}$), and a modest negative correlation with CRF-slope (r_g=-0.14, se=0.06, p=$1.0x10^{-2}$). These two genetic correlations suggest that higher levels of activity have a shared genetic contribution with greater fitness, and with a lower rate of increase of heart rate

during exercise. The within-sex between-phenotype genetic correlations were similar to those for the sex-combined results (Additional file 1: **Figure S41**). Within-phenotype between-sex genetic correlations were all near 1.0, reflecting that the overall genomic signal is highly correlated despite a few different loci between sexes (Additional file 1: **Figure S41**).

Genetic correlations with multiple traits and diseases

A comprehensive analysis of genetic correlations of CRF-vo2max, CRF-slope and PA with a wide array of traits was performed using LD-HUB (**Figure 5** shows the combined-sex result, Additional file 1: **Figures S42-S44** show by-sex results). All genetic correlations described below were FDR-significant in the sex-combined sample. Additional file 1: **Table S9** includes the sex-combined, male, and female FDR-significant genetic correlations, as well as heritabilities, associated standard errors, and GWAS study PubMed IDs for all FDR-significant traits and diseases.

Shared genetic risk: lipids, glycaemic and cardiometabolic traits

Genetic variation associated with lower CRF-vo2max was correlated with features of the typical type-2 diabetes (T2D) profile: higher insulin resistance (HOMA-IR: r_g=-0.58, se=0.12) and higher fasting insulin (r_g=-0.45, se=0.09), impaired beta-cell function (HOMA-B: r_g=-0.40, se=0.09), hyperglycaemia (HbA1c: r_g=-0.25, se=0.08) and hypertriglyceridemia (r_g=-0.17, se=0.04). In the PA combined sample, as well as genetic correlations with fasting insulin (r_g=-0.31, se=0.07), insulin resistance (r_g=-0.29, se=0.07), and triglycerides (r_g=-0.20, se=0.04), we observed a positive correlation with HDL cholesterol (r_g=0.28, se=0.04) and a direct negative correlation with T2D (r_g=-0.23, se=0.05). Genetic variation associated with lower levels of physical activity was associated with increased risk of T2D.

Both CRF-vo2max and PA were negatively correlated with coronary artery disease (CRF-vo2max r_g=-0.15, se=0.05; PA r_g=-0.21, se=0.04). Like T2D, genetic variation associated with lower CRF-vo2max and lower PA was associated with increased risk for coronary artery disease. Additional file 1: **Figure S45** highlights the FDR-significant glycaemic, cardiometabolic, and lipid traits in the combined sample, as well as in males and females separately.

Shared genetic risk: metabolites, hormones

CRF-vo2max was significantly genetically correlated with glycoprotein acetyls, GlycA (r_g=-029, se=0.10), such that genetic variation associated with greater CRF-vo2max was associated with lower levels of the GlycA biomarker. Elevated blood concentrations of GlycA are a strong predictor for long-term risk of morbidity and mortality from diverse diseases (74).

For PA we observed positive FDR-significant genetic correlations with several related HDL metabolites. The largest HDL metabolite correlation with PA was total cholesterol in medium HDL (r_g=0.42, se=0.15). Genetic

variation associated with greater CRF-vo2max and PA were associated with higher levels of HDL. PA was also significantly correlated with acetate (r_g=0.36, se=0.13) and apolipoprotein A-I (r_g=0.35, se=0.12) – the major protein component of HDL which enables fat efflux from within cells for transport to the liver for excretion. Additional file 1: **Figure S50** shows the genetic correlations with blood metabolites for the combined sample, and males and females separately.

Circulating levels of the satiety hormone leptin were negatively correlated with CRF-vo2max and PA (CRF-vo2max r_g=-0.41, se=0.08; PA r_g=-0.37, se=0.07), and for BMI-adjusted leptin (CRF-vo2max r_g=-0.32, se=0.08; PA r_g=-0.21, se=0.07) (Additional file 1: **Figure S45**). The genetic profile associated with greater levels of fitness and activity is correlated with greater circulating levels of leptin, which inhibits hunger and reduces fat storage in adipocytes.

Shared genetic risk: body measures

CRF-vo2max was positively genetically correlated with height (r_g=0.13, se=0.04), birth weight (r_g=0.14, se=0.05) and own birth weight (r_g=0.13, se=0.04). All other FDR-significant anthropometric correlations were negative (r_g range=-0.23 – -0.44, se range=0.04 – 0.07): body fat, waist circumference, measures of obesity, BMI, waist-to-hip ratio and hip circumference. CRF-slope had FDR-significant negative genetic correlations with childhood obesity (r_g=-0.24, se=0.06) and childhood height (r_g=-0.26, se=0.07), BMI (r_g=-0.18, se=0.05), birth weight (r_g=-0.15, se=0.05) and own birth weight (r_g=-0.12, se=0.04). For PA, all FDR-significant body measures including height were negative (r_g=-0.12 – -0.37, se=0.03 – 0.06): body fat, waist circumference, measures of obesity, BMI, waist-to-hip ratio and hip circumference, and height. Genetic variation associated with greater CRF-vo2max and higher levels of PA was associated with lower body fat, BMI, obesity, and waist and hip circumferences. Additional file 1: **Figures S46**–**S48** show FDR-significant genetic correlations with anthropometric measures for the combined sample and males and females separately.

Shared genetic risk: lung function and heart rate

CRF-vo2max was positively associated with lung function: forced vital capacity (FVC: r_g=0.22 – 0.26, se=0.03 – 0.06), forced expiratory volume in 1 second (FEV1: r_g=0.18 – 0.24, se=0.03 – 0.07), and peak expiratory flow (r_g=0.12, se=0.03). CRF-slope was negatively correlated with FVC (r_g=-0.13, se=0.04) and peak expiratory flow (r_g=-0.14, se=0.04). For PA, we found modest positive genetic correlations with FVC (r_g=0.13 – 0.15, se=0.03 – 0.04) and FEV1 (r_g=0.10, se=0.02). Genetic variation associated with both fitness and activity is associated with improved lung function.

Only CRF-slope was genetically correlated with resting heart rate (r_g=-0.31, se=0.07). Genetic correlation associated with greater acceleration of heart rate during exercise was associated a lower resting heart rate. Additional file 1: **Figure S49** shows genetic correlations with lung function and heart rate, for the combined sample and males and females separately.

Shared genetic risk: smoking behaviour and longevity

We observed a positive genetic correlation between PA and former vs current smoking status (r_g=0.27, se=0.07). A genotype associated with greater levels of activity associated with smoking cessation.

Finally, we found that genetic variation associated with greater CRF-vo2max and PA was also significantly associated with longevity, as measured by father's age at death (CRF-vo2max r_g=0.37, se=0.09; PA r_g=0.30, se=0.07) and parents' ages at death (PA r_g=0.28, se=0.08). FDR-significant genetic correlations with smoking behaviour, cancer and longevity reported above are highlighted in Additional file 1: **Figure S51**.

Shared genetic risk: demographics

CRF-vo2max and PA were positively correlated with years of schooling (CRF-vo2max r_g=0.28 – 0.32, se=0.04 – 0.05; PA r_g=0.11 – 0.14, se=0.03 – 0.04) and age at birth of first child (CRF-vo2max r_g=0.22, se=0.05; PA r_g=0.13, se=0.04). CRF-slope had a corresponding negative genetic correlation with years of schooling (r_g=-0.15 – -0.17, se=0.05). FDR-significant genetic correlations with education and reproductive phenotypes are highlighted in Additional file 1: **Figure S52**.

Discussion

Genetic variation explains a modest but significant proportion of the variation in vo2max, rate of increase of heart rate during exercise, and average weekly activity level, for both sexes. Genetic association signals were highly correlated between sexes, with only six SNPs that showed evidence for a sex-specific association (CRF-vo2max: rs1741294 male; CRF-slope: rs111299422 female, rs1741294 male; PA rs75986475, rs13329850 female, rs78661713 male). The two derived fitness traits, CRF-vo2max (vo2max) and CRF-slope (rate of increase of heart rate during exercise), were negatively correlated, phenotypically and genetically. Greater levels of fitness corresponded to a slower rate of increase of heart rate during exercise. CRF-vo2max was positively correlated with PA; greater fitness was associated with higher levels of physical activity.

Lower CAD and T2D among most fit and active

We found that higher levels of CRF-vo2max and PA were associated with better health outcomes in a dose-dependent way. The frequency of CAD and diabetes mellitus diagnosis was substantially lower among individuals with higher levels of CRF-vo2max and PA, even after controlling for BMI and SES. In this large UK population-based sample, we find the same associations between PA and CRF-vo2max observed in epidemiological and clinical studies of coronary heart disease and cardiovascular events (75) and T2D (76). As has been previously observed (3), there was a noteworthy steeper drop-off in disease occurrence with

increased PA and CRF-vo2max among the least fit and active third of the sample. The fitness and activity associations with disease observed in the combined sample were most stark among males. For CRF-vo2max and CRF-slope, disease diagnosis frequency was generally lower as we only included individuals with no contraindications for the standard cycle ramp test protocol. Disease diagnosis association with CRF-slope was noisy, with no clear pattern of association except for a small reduction in CAD and diabetes mellitus. It is not clear why there would be a reduction in disease diagnoses with increasing CRF-slope, given the negative correlation with CRF-vo2max. One possible factor is that among the participants with high CRF-slope levels there is a mix of the least fit samples whose heart rates increased by a larger amount during exercise, participants with lower resting heart rates due to medical conditions and medication (expected to be low given risk category exclusions), and the most fit samples who would have lower resting heart rates.

Disease associated genes implicated in CRF and PA

Our functional genetic analyses identified genes associated with CRF and PA, with known associations in disease and clinical vital sign related phenotypes. For CRF-vo2max, we found functional candidates associated with expression in heart, artery, lung, brain, adipose, and skeletal muscle. All functional candidates were within 17q21.31, with five genes also identified in the gene-based association analysis (*WNT3*, *MAPT*, *LRRC37A2*, *LRRC37A*, *CRHR1*). These five genes have previously been associated with GWASs for atrial fibrillation (46), diabetes mellitus, and coronary artery disease (47), Parkinson's (48), and Alzheimer's disease (49), lung function (FEV1, FVC (50), peak expiratory flow (51)), heart and electrocardiography measures (52), cardiac arrhythmia (53), and blood phenotypes (systolic blood pressure, eosinophil count (54), haemoglobin level (55), haematocrit (56), red blood cell density (57)). Subsets of CRF-vo2max functional candidates were enriched in the GWAS catalog for multiple system atrophy, Alzheimer's disease, Parkinson's disease, idiopathic pulmonary fibrosis, primary biliary cirrhosis, FEV1, haematocrit, and haemoglobin concentration.

For CRF-slope, functional candidates fell within one of two cytogenic regions, 7q21.1 and 8p23.1. *ACHE* within 7q21.1, also identified as a positional candidate 52.87kb upstream of rs4582488, was significantly associated with expression in adipose, artery and heart tissue, and has previously been associated with resting heart rate (58), heart rate response to exercise and recovery (59), electrocardiography (60), diastolic blood pressure (61), and diabetes mellitus (62). The gene-based association identified genes (*TRIP6*, *PINX1*, and *ERI1*) were significantly associated with expression change in whole blood and lung, and have previously been associated with blood pressure (63), triglyceride levels (64), HDL cholesterol (65), diabetes mellitus (66), and atherosclerosis (67). Other promising functional candidates (*SRRT*, *RP11-62H7.2*, *FAM86B3P*, *FAM85B*, and *AF131215.2*) were associated with expression in lung, heart, skeletal muscle, and coronary and tibial artery, and have been associated in previous GWASs with heart rate response to exercise and recovery (59), electrocardiography (68), diabetes mellitus (66), and obesity (69). We found subsets of the CRF-slope functional candidates were enriched in GWAS gene sets for the thrombosis and atherosclerosis risk factor *PAI-1*, heart rate response to exercise and recovery, and resting heart rate.

For PA, the functional candidates on 17q21.31 (*MAPT*, *LRRC37A2*, *LRRC37A*, *KANSL1*, *CRHR1*, *ARL17A*, and *ARHGAP27*) and 22q13.2 (*L3MBTL2*) with a high probability of the same variant driving PA association and expression in adipose, brain, heart, lung and muscle, have previously been associated with disease (Parkinson's disease (70), hypertrophic cardio myopathy (71), ischaemic stroke, diabetes mellitus and coronary artery disease (72)) anthropometrics (BMI-adjusted waist hip to ratio, BMI (73)), blood phenotypes (haemoglobin measurement (55), red blood cell density, hematocrit (57), eosinophil count and systolic blood pressure (54)), lung function (FVC (54), FEV1 (51)), and heart and electrocardiography measures (QRS measures (52), cardiac arrhythmia (53)). We found PA functional candidates enriched in Parkinson's and Alzheimer's disease, noteworthy given their underlying insulin resistance pathology (although the insulin resistance link is controversial for Alzheimer's disease), idiopathic pulmonary fibrosis, multiple system atrophy and primary biliary cirrhosis. Functional candidates from the female PA GWAS were also enriched in the lung function phenotypes FEV1 and FVC.

Genetic correlations link CRF and PA with disease

Genetic correlations complement epidemiological and clinical observations. Because inherited genetic risk cannot be due to confounding, genetic correlations limit the number of potential confounders linked through genotype. The shared genetic architecture of both CRF-vo2max and PA clearly implicated the lifestyle-related chronic diseases obesity, T2D, and CAD, and their body composition and metabolic risk factors. CRF-vo2max and PA were negatively genetically correlated with obesity, body fat, body mass index, and waist-to-hip circumference, triglycerides, the satiety hormone leptin, fasting insulin, insulin resistance, and T2D (significant only for PA), and CAD. These negative genetic correlations corroborate the epidemiological associations, previous research findings, and clinical observations of CRF and PA and body composition (77, 78), lipid profile (79, 80), adipocytokines (81), fasting insulin and insulin resistance (82-86) and leptin (76), T2D (87-89), and CAD (90). Likewise, the positive genetic correlations are consistent with previous observations of positive relationships between CRF and PA, and HDL cholesterol (80, 91) and longevity (92-94).

CRF-vo2max was negatively genetically correlated with blood levels of glycoprotein acetyls (GlycA). GlycA is associated with both acute and chronic inflammation and is a strong predictor for long-term risk of morbidity and mortality from a wide range of diseases (74).

Additionally, CRF-vo2max and PA had positive genetic correlations with age at birth of first child and level of education, respectively a correlate and component of SES. CRF-slope showed a corresponding negative correlation with years of education. This is consistent with population-based observations of a socioeconomic gradient in PA and CRF: people living in higher socioeconomic status areas tend to be more active and more fit, and from a public health perspective particularly important as "reduced PA and fitness levels may contribute to social inequalities in health" (95).

We note several limitations. The gold standard for CRF measurement involves exercise to volitional exhaustion with measurement of gas exchange (uptake of oxygen, elimination of carbon dioxide). A submaximal CRF test with extrapolation of VO2max scores is however a much more practical solution, especially for the collection of a large sample in the age ranges studied. The objective PA monitoring only captured a 7-day window, effectively a snapshot of participant behaviour, and we do not know if this reflects their habitual behaviour. We also used raw acceleration and did not attempt to tease apart PA intensity, recognized as important for health outcomes. All genetic associations, especially the sex-specific SNPs identified, will require replication in an independent sample. In our genetic correlation analyses, some of the largest correlations were observed with blood metabolites but these were based on relatively small GWAS reference samples (metabolites N=24,925 compared to body fat N=100,716) and many metabolite correlations failed to pass the multiple testing threshold. As well as this sample size (or power of each discovery GWAS) limitation, genetic correlations cannot resolve reverse causation. Another limitation is correlated lifestyle risk factors not considered, in particular eating behaviours and nutrition. Eating habits, diet, and exercise behaviours tend to go hand in hand, in large part influenced by the environment, e.g., sedentary time spent watching television and eating high-fat foods (96, 97). A holistic approach to health – addressing eating habits and including an effective dietary intervention (98), alongside prescription PA and monitoring of CRF – may be the most effective practical approach, especially in already overweight individuals given the possibility that there exists a bi-directional relationship between adiposity and PA (99). Despite the complex interplay of eating habits, nutrition, environment, and PA, CRF as a regularly measured vital sign is an objective and easily measured sentinel for health.

Conclusions

In summary, we found strong evidence that genetic variants associated with CRF and PA also influenced genetic expression in a small set of genes in heart, artery, brain, lung, muscle, and adipose tissue. These functionally relevant genes were enriched among genes known to be associated with CAD, T2D and Alzheimer's disease – three of the top 10 causes of death in high-income countries – as well as Parkinson's, pulmonary fibrosis, and blood pressure, heart rate, and respiratory phenotypes. Finally, genetic correlation related lower fitness and activity with several disease risk factors including greater waist-to-hip ratio, BMI, and obesity; with a typical T2D profile including higher insulin resistance, higher fasting glucose, impaired beta-cell function, hyperglycaemia, hypertriglyceridemia; increased risk for CAD and T2D; and shorter lifespan. Genetics supports three decades of evidence for the inclusion of CRF as a clinical vital sign (3). CRF is a window onto general health and wellbeing, and physical activity is the primary means of improving CRF. In fact, to meet the WHO policy of reduction in sedentary lifestyles that carry a heavy cost to personal and public health, "policies to increase population levels of physical activity need to be prioritised and scaled up urgently" (9). PA too is "vital"; we think both measures provide important information and ideally should be collected in tandem if feasible. Given the global health burden of non-communicable disease related to lifestyle choices, as well as dietary advice, regular measurement of CRF as a marker of health and routine prescription of PA is imperative.

Figure captions

Figure 1. Heart rate signal processing for a single random sample. Trend = repeated measurements within the 'Exercise' phase of the submaximal cycle ramp test; Raw = unaltered heart rate data; Butterworth filtered = heart rate data after low-pass Butterworth filter (heart rate at the beginning and end of the exercise period on this filtered signal used to calculate CRF-vo2max); Linear model = fitted heart rate data values from a linear model fitted to the Butterworth filtered signal (CRF-slope is the slope of this linear model).

Figure 2. ICD-10 diagnosis frequency by CRF and PA level. Frequency in the UKB for the top 10 global causes of death in high-income countries in 2016, as reported by the World Health Organization (WHO, Additional file 1: **Table S3**). The CRF and PA phenotypes above were controlled for age and socioeconomic status (SES, the Townsend deprivation index) – they are the residuals from a regression on age and SES. For some diseases, e.g., breast cancer, disease trends are only meaningful separated by sex. CAD = ischaemic heart disease; STR = stroke; ALZ = Alzheimer's disease and other dementias; LCX = trachea, bronchus and lung cancers; COPD = chronic obstructive pulmonary disease; LRTI = lower respiratory tract infections; CCX = colon and rectum cancers; DIA = diabetes mellitus; KID = kidney diseases; BCX = breast cancers

Figure 3. Genome-wide association signal by sex for CRF-vo2max, CRF-slope, and PA. The red line shows genome-wide significance, $p=5x10^{-8}$.

Figure 4. TWAS and colocalization identified genes. Genes listed on the y-axes are significant at the conservative Bonferroni corrected TWAS p-value. On the x-axis is tissue type; category of tissue is represented by colour, colocalization posterior probability 4 (COLOC PP4 – the probability of a shared causal SNP driving the association signal and expression in a reference tissue) represented by line angle (vertical line is a probability of 1.0 that the GWAS-identified SNP is an eQTL for the gene). The complementary TWAS and colocalization approaches provide a priority list of genes for further exploration. Genes in bold were significant in the gene-based test of association.

Figure 5. Genetic correlations for combined males and females. red = negative correlations; blue = positive correlations; size = relative size of the genetic correlation; Bonferroni threshold shows p=0.05/total number of correlations per trait; False discovery rate threshold indicates the largest FDR-adjusted p-value < 0.05

List of abbreviations

CRF = cardiorespiratory fitness, PA = physical activity, AHA = American Heart Association, UKB = UK Biobank, CRF-vo2max = an estimate of the participants' maximum volume of oxygen uptake, per kilogram of body weight, per minute, CRF-slope = an estimate of the rate of increase of heart rate during exercise, GWAS = genome-wide association study, CAD = coronary artery disease/ ischaemic heart disease, T2D = type 2 diabetes, h^2 = heritability, h^2_{SNP} = SNP heritability, ECG = electrocardiograph, BMI = body mass index, HR = heart rate, WL = workload, WHO = world health organization, ICD-10 = International Classification of Diseases revision 10, SES = socioeconomic status, TWAS = transcriptome-wide association, FDR = false-discovery rate, eQTL = expression quantitative trait locus, PP4 = posterior probability 4

Declarations

Ethics approval and consent to participate

Ethics approval for the UK Biobank study was obtained from the North West Centre for Research Ethics Committee (11/NW/0382); participants provided electronic signed consent at recruitment. This research has been conducted under application number 13427.

Consent for publication

Not applicable.

Availability of data and materials

The data that support the findings of this study are available upon application to the UK Biobank (UKB). The UKB database is globally accessible to approved researchers.

Competing interests

The authors declare that they have no competing interests.

Funding

This research was funded by the National Institute for Health Research (NIHR) Biomedical Research Centre (BRC) South London and Maudsley, and Guy's and St Thomas', NHS Foundation Trusts, and King's College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. CML receives funding from a UK Medical Research Council (MRC) grant (MR/N015746/1).

Authors' contributions

KBH, MH, CML conceived and designed the study. KBH analysed the data. KBH, EP, MT, KPG, MH, JRIC, CML interpreted the results and wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We would like to acknowledge the UKB participants for their contributions.

Additional files

Additional file 1

Additional file 1.docx includes supplementary Methods (genetic data quality control, GWAS considerations, alternate genetic association models, and post GWAS details), Results (including alternate association models, GWAS post-processing), Tables S1 – S9, and Figures S1 – S52.

References

1. DeFina LF, Haskell WL, Willis BL, Barlow CE, Finley CE, Levine BD, et al. Physical activity versus cardiorespiratory fitness: two (partly) distinct components of cardiovascular health? Prog Cardiovasc Dis. 2015;57(4):324-9.

2. Lavie CJ, Johannsen N, Swift D, Senechal M, Earnest C, Church T, et al. Exercise is Medicine - The Importance of Physical Activity, Exercise Training, Cardiorespiratory Fitness and Obesity in the Prevention and Treatment of Type 2 Diabetes. Eur Endocrinol. 2014;10(1):18-22.

3. Ross R, Blair SN, Arena R, Church TS, Despres JP, Franklin BA, et al. Importance of Assessing Cardiorespiratory Fitness in Clinical Practice: A Case for Fitness as a Clinical Vital Sign: A Scientific Statement From the American Heart Association. Circulation. 2016;134(24):e653-e99.

4. Stamatakis E, Hamer M, O'Donovan G, Batty GD, Kivimaki M. A non-exercise testing method for estimating cardiorespiratory fitness: associations with all-cause and cardiovascular mortality in a pooled analysis of eight population-based cohorts. Eur Heart J. 2013;34(10):750-8.

5. Tipton CM. The history of "Exercise Is Medicine" in ancient civilizations. Adv Physiol Educ. 2014;38(2):109-17.

826 6. Mokdad AH, Marks JS, Stroup DF, Gerberding JL. Actual causes of death in the United States, 2000.
827 JAMA. 2004;291(10):1238-45.

828 7. Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT, et al. Effect of physical inactivity
829 on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy.
830 Lancet. 2012;380(9838):219-29.

831 8. Ding D, Lawson KD, Kolbe-Alexander TL, Finkelstein EA, Katzmarzyk PT, van Mechelen W, et al. The
832 economic burden of physical inactivity: a global analysis of major non-communicable diseases. Lancet.
833 2016;388(10051):1311-24.

834 9. Guthold R, Stevens GA, Riley LM, Bull FC. Worldwide trends in insufficient physical activity from
835 2001 to 2016: a pooled analysis of 358 population-based surveys with 1.9 million participants. Lancet Glob
836 Health. 2018;6(10):e1077-e86.

837 10. British Heart Foundation Physical Inactivity Report 2017 [Available from:
838 https://www.bhf.org.uk/informationsupport/publications/statistics/physical-inactivity-report-2017]

839 11. Teran-Garcia M, Rankinen T, Bouchard C. Genes, exercise, growth, and the sedentary, obese child.
840 J Appl Physiol (1985). 2008;105(3):988-1001.

841 12. den Hoed M, Brage S, Zhao JH, Westgate K, Nessa A, Ekelund U, et al. Heritability of objectively
842 assessed daily physical activity and sedentary behavior. Am J Clin Nutr. 2013;98(5):1317-25.

843 13. Timmons JA, Knudsen S, Rankinen T, Koch LG, Sarzynski M, Jensen T, et al. Using molecular
844 classification to predict gains in maximal aerobic capacity following endurance exercise training in humans.
845 J Appl Physiol (1985). 2010;108(6):1487-96.

846 14. Bouchard C, Sarzynski MA, Rice TK, Kraus WE, Church TS, Sung YJ, et al. Genomic predictors of the
847 maximal O(2) uptake response to standardized exercise training programs. J Appl Physiol (1985).
848 2011;110(5):1160-70.

849 15. Williams CJ, Williams MG, Eynon N, Ashton KJ, Little JP, Wisloff U, et al. Genes to predict VO2max
850 trainability: a systematic review. BMC Genomics. 2017;18(Suppl 8):831.

851 16. Klimentidis YC, Raichlen DA, Bea J, Garcia DO, Wineinger NE, Mandarino LJ, et al. Genome-wide
852 association study of habitual physical activity in over 377,000 UK Biobank participants identifies multiple
853 variants including CADM2 and APOE. Int J Obes (Lond). 2018.

854 17. Doherty A, Smith-Bryne K, Ferreira T, Holmes MV, Holmes C, Pulit S, et al. GWAS identifies 14 loci
855 for device-measured physical activity and sleep duration. Nature Communications. 2018;9(5257).

856 18. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with
857 deep phenotyping and genomic data. Nature. 2018;562(7726):203-9.

858 19. Wilson MG, Ellison GM, Cable NT. Basic science behind the cardiovascular benefits of exercise. Brit
859 J Sport Med. 2016;50(2):93-+.

860 20. Karastergiou K, Smith SR, Greenberg AS, Fried SK. Sex differences in human adipose tissues - the
861 biology of pear shape. Biol Sex Differ. 2012;3.

862 21. Valencak TG, Osterrieder A, Schulz TJ. Sex matters: The effects of biological sex on adipose tissue
863 biology and energy metabolism. Redox Biol. 2017;12:806-13.

864 22. van Uffelen JGZ, Khan A, Burton NW. Gender differences in physical activity motivators and
865 context preferences: a population-based study in people in their sixties. BMC Public Health.
866 2017;17(1):624.

867 23. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access
868 resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med.
869 2015;12(3):e1001779.

870 24. UK Biobank Cardio Assessment Version 1.0 [Available from:
871 https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/Cardio.pdf]

872 25. Celis-Morales CA, Lyall DM, Anderson J, Iliodromiti S, Fan Y, Ntuk UE, et al. The association
873 between physical activity and risk of mortality is modulated by grip strength and cardiorespiratory fitness:
874 evidence from 498 135 UK-Biobank participants. Eur Heart J. 2017;38(2):116-22.

875 26. Tanaka H, Monahan KD, Seals DR. Age-predicted maximal heart rate revisited. J Am Coll Cardiol.
876 2001;37(1):153-6.

877 27. Swain DP. Energy cost calculations for exercise prescription: an update. Sports Med. 2000;30(1):17-
878 22.

879 28. Kim Y, White T, Wijndaele K, Westgate K, Sharp SJ, Helge JW, et al. The combination of
880 cardiorespiratory fitness and muscle strength, and mortality risk. Eur J Epidemiol. 2018;33(10):953-64.

881 29. Doherty A, Jackson D, Hammerla N, Plotz T, Olivier P, Granat MH, et al. Large Scale Population
882 Assessment of Physical Activity Using Wrist Worn Accelerometers: The UK Biobank Study. PLoS One.
883 2017;12(2):e0169649.

884 30. UK Biobank Resource 664 [Available from: http://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=664.]

885 31. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical
886 Computing, Vienna, Austria. https://www.R-project.org/.

887 32. Van Rossum G, Drake FL. Python 3 Reference Manual. Scotts Valley, CA: CreateSpace; 2009.

888 33. Koster J, Rahmann S. Snakemake--a scalable bioinformatics workflow engine. Bioinformatics.
889 2012;28(19):2520-2.

890 34. Hanscombe KB, Coleman RI, Traylor M, Lewis CM. ukbtools: An R package to manage and query UK
891 Biobank data. PLoS One. 2019.

892 35. Abraham G, Qiu Y, Inouye M. FlashPCA2: principal component analysis of Biobank-scale genotype
893 datasets. Bioinformatics. 2017;33(17):2776-8.

894 36. Loh PR, Tucker G, Bulik-Sullivan BK, Vilhjalmsson BJ, Finucane HK, Salem RM, et al. Efficient
895 Bayesian mixed-model analysis increases association power in large cohorts. Nat Genet. 2015;47(3):284-90.

896 37. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to
897 the challenge of larger and richer datasets. Gigascience. 2015;4:7.

898 38. Bulik-Sullivan BK, Loh P-R, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the
899 Psychiatric Genomics C, et al. LD Score regression distinguishes confounding from polygenicity in genome-
900 wide association studies. Nature Genetics. 2015;47:291.

901 39. Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BW, et al. Integrative approaches for large-scale
902 transcriptome-wide association studies. Nat Genet. 2016;48(3):245-52.

903 40. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of
904 genetic associations with FUMA. Nat Commun. 2017;8(1):1826.

905 41. Zheng J, Erzurumluoglu AM, Elsworth BL, Kemp JP, Howe L, Haycock PC, et al. LD Hub: a
906 centralized database and web interface to perform LD score regression that maximizes the potential of
907 summary level GWAS data for SNP heritability and genetic correlation analysis. Bioinformatics.
908 2017;33(2):272-9.

42. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics. 2015;31(21):3555-7.

43. Aschard H, Vilhjalmsson BJ, Joshi AD, Price AL, Kraft P. Adjusting for heritable covariates can bias effect estimates in genome-wide association studies. Am J Hum Genet. 2015;96(2):329-39.

44. Stefansson H, Helgason A, Thorleifsson G, Steinthorsdottir V, Masson G, Barnard J, et al. A common inversion under selection in Europeans. Nat Genet. 2005;37(2):129-37.

45. Permuth-Wey J, Lawrenson K, Shen HC, Velkova A, Tyrer JP, Chen Z, et al. Identification and molecular characterization of a new ovarian cancer susceptibility locus at 17q21.31. Nat Commun. 2013;4:1627.

46. Nielsen JB, Thorolfsdottir RB, Fritsche LG, Zhou W, Skov MW, Graham SE, et al. Biobank-driven genomic discovery yields new insight into atrial fibrillation biology. Nat Genet. 2018;50(9):1234-9.

47. van der Harst P, Verweij N. Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. Circ Res. 2018;122(3):433-43.

48. Pankratz N, Beecham GW, DeStefano AL, Dawson TM, Doheny KF, Factor SA, et al. Meta-analysis of Parkinson's disease: identification of a novel locus, RIT2. Ann Neurol. 2012;71(3):370-84.

49. Jun G, Ibrahim-Verbaas CA, Vronskaya M, Lambert JC, Chung J, Naj AC, et al. A novel Alzheimer disease locus located near the gene encoding tau protein. Mol Psychiatry. 2016;21(1):108-17.

50. Wyss AB, Sofer T, Lee MK, Terzikhan N, Nguyen JN, Lahousse L, et al. Multiethnic meta-analysis identifies ancestry-specific and cross-ancestry loci for pulmonary function. Nat Commun. 2018;9(1):2976.

51. Shrine N, Guyatt AL, Erzurumluoglu AM, Jackson VE, Hobbs BD, Melbourne CA, et al. New genetic signals for lung function highlight pathways and chronic obstructive pulmonary disease associations across multiple ancestries. Nat Genet. 2019;51(3):481-93.

52. van der Harst P, van Setten J, Verweij N, Vogler G, Franke L, Maurano MT, et al. 52 Genetic Loci Influencing Myocardial Mass. J Am Coll Cardiol. 2016;68(13):1435-48.

53. Ishigaki K, Akiyama M, Kanai M, Takahashi A, Kawakami E, Sugishita H, et al. Large-scale genome-wide association study in a Japanese population identifies novel susceptibility loci across different diseases. Nat Genet. 2020;52(7):669-79.

937 54. Kichaev G, Bhatia G, Loh PR, Gazal S, Burch K, Freund MK, et al. Leveraging Polygenic Functional
938 Enrichment to Improve GWAS Power. Am J Hum Genet. 2019;104(1):65-75.

939 55. Oskarsson GR, Oddsson A, Magnusson MK, Kristjansson RP, Halldorsson GH, Ferkingstad E, et al.
940 Predicted loss and gain of function mutations in ACO1 are associated with erythropoiesis. Commun Biol.
941 2020;3(1):189.

942 56. Astle WJ, Elding H, Jiang T, Allen D, Ruklisa D, Mann AL, et al. The Allelic Landscape of Human
943 Blood Cell Trait Variation and Links to Common Complex Disease. Cell. 2016;167(5):1415-29 e19.

944 57. Chen MH, Raffield LM, Mousas A, Sakaue S, Huffman JE, Moscati A, et al. Trans-ethnic and
945 Ancestry-Specific Blood-Cell Genetics in 746,667 Individuals from 5 Global Populations. Cell.
946 2020;182(5):1198-213 e14.

947 58. Ramirez J, Duijvenboden SV, Ntalla I, Mifsud B, Warren HR, Tzanis E, et al. Thirty loci identified for
948 heart rate response to exercise and recovery implicate autonomic nervous system. Nat Commun.
949 2018;9(1):1947.

950 59. Verweij N, van de Vegte YJ, van der Harst P. Genetic study links components of the autonomous
951 nervous system to heart-rate profile during exercise. Nat Commun. 2018;9(1):898.

952 60. Verweij N, Benjamins JW, Morley MP, van de Vegte YJ, Teumer A, Trenkwalder T, et al. The
953 Genetic Makeup of the Electrocardiogram. Cell Syst. 2020;11(3):229-38 e5.

954 61. Giri A, Hellwege JN, Keaton JM, Park J, Qiu C, Warren HR, et al. Trans-ethnic association study of
955 blood pressure determinants in over 750,000 individuals. Nat Genet. 2019;51(1):51-62.

956 62. Sim X, Ong RT, Suo C, Tay WT, Liu J, Ng DP, et al. Transferability of type 2 diabetes implicated loci
957 in multi-ethnic cohorts from Southeast Asia. PLoS Genet. 2011;7(4):e1001363.

958 63. Warren HR, Evangelou E, Cabrera CP, Gao H, Ren M, Mifsud B, et al. Genome-wide association
959 analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. Nat
960 Genet. 2017;49(3):403-15.

961 64. Ripatti P, Ramo JT, Mars NJ, Fu Y, Lin J, Soderlund S, et al. Polygenic Hyperlipidemias and
962 Coronary Artery Disease Risk. Circ Genom Precis Med. 2020;13(2):e002725.

963 65. Richardson TG, Sanderson E, Palmer TM, Ala-Korpela M, Ference BA, Davey Smith G, et al.
964 Evaluating the relationship between circulating lipoprotein lipids and apolipoproteins with risk of coronary
965 heart disease: A multivariable Mendelian randomisation analysis. PLoS Med. 2020;17(3):e1003062.

966 66. Xue A, Wu Y, Zhu Z, Zhang F, Kemper KE, Zheng Z, et al. Genome-wide association analyses
967 identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. Nat Commun.
968 2018;9(1):2941.

969 67. Bis JC, Kavousi M, Franceschini N, Isaacs A, Abecasis GR, Schminke U, et al. Meta-analysis of
970 genome-wide association studies from the CHARGE consortium identifies common variants associated with
971 carotid intima media thickness and plaque. Nat Genet. 2011;43(10):940-7.

972 68. van Setten J, Verweij N, Mbarek H, Niemeijer MN, Trompet S, Arking DE, et al. Genome-wide
973 association meta-analysis of 30,000 samples identifies seven novel loci for quantitative ECG traits. Eur J
974 Hum Genet. 2019;27(6):952-62.

975 69. Schlauch KA, Read RW, Lombardi VC, Elhanan G, Metcalf WJ, Slonim AD, et al. A Comprehensive
976 Genome-Wide and Phenome-Wide Examination of BMI and Obesity in a Northern Nevadan Cohort. G3
977 (Bethesda). 2020;10(2):645-64.

978 70. Edwards TL, Scott WK, Almonte C, Burt A, Powell EH, Beecham GW, et al. Genome-wide association
979 study confirms SNPs in SNCA and the MAPT region as common risk factors for Parkinson disease. Ann Hum
980 Genet. 2010;74(2):97-109.

981 71. Tadros R, Francis C, Xu X, Vermeer AMC, Harper AR, Huurman R, et al. Shared genetic pathways
982 contribute to risk of hypertrophic and dilated cardiomyopathies with opposite directions of effect. Nat
983 Genet. 2021;53(2):128-34.

984 72. Song Y, Choi JE, Kwon YJ, Chang HJ, Kim JO, Park DH, et al. Identification of susceptibility loci for
985 cardiovascular disease in adults with hypertension, diabetes, and dyslipidemia. J Transl Med.
986 2021;19(1):85.

987 73. Zhu Z, Guo Y, Shi H, Liu CL, Panganiban RA, Chung W, et al. Shared genetic and experimental links
988 between obesity-related traits and asthma subtypes in UK Biobank. J Allergy Clin Immunol.
989 2020;145(2):537-49.

990 74. Kettunen J, Ritchie SC, Anufrieva O, Lyytikainen LP, Hernesniemi J, Karhunen PJ, et al. Biomarker
991 Glycoprotein Acetyls Is Associated With the Risk of a Wide Spectrum of Incident Diseases and Stratifies
992 Mortality Risk in Angiography Patients. Circ Genom Precis Med. 2018;11(11):e002234.

993 75. Kodama S, Saito K, Tanaka S, Maki M, Yachi Y, Asumi M, et al. Cardiorespiratory fitness as a
994 quantitative predictor of all-cause mortality and cardiovascular events in healthy men and women: a meta-
995 analysis. JAMA. 2009;301(19):2024-35.

996 76. Lin X, Zhang X, Guo J, Roberts CK, McKenzie S, Wu WC, et al. Effects of Exercise Training on
997 Cardiorespiratory Fitness and Biomarkers of Cardiometabolic Health: A Systematic Review and Meta-
998 Analysis of Randomized Controlled Trials. J Am Heart Assoc. 2015;4(7).

999 77. L. dL, R. R. Physical Activity, Cardiorespiratory Fitness, and Obesity. In: P. K, P. N, editors.
1000 Cardiorespiratory Fitness in Cardiometabolic Diseases: Springer; 2019.

1001 78. Prioreschi A, Brage S, Westgate K, Norris SA, Micklesfield LK. Cardiorespiratory fitness levels and
1002 associations with physical activity and body composition in young South African adults from Soweto. BMC
1003 Public Health. 2017;17(1):301.

1004 79. Sui X, Sarzynski MA, Lee DC, Kokkinos PF. Impact of Changes in Cardiorespiratory Fitness on
1005 Hypertension, Dyslipidemia and Survival: An Overview of the Epidemiological Evidence. Prog Cardiovasc Dis.
1006 2017;60(1):56-66.

1007 80. Breneman CB, Polinski K, Sarzynski MA, Lavie CJ, Kokkinos PF, Ahmed A, et al. The Impact of
1008 Cardiorespiratory Fitness Levels on the Risk of Developing Atherogenic Dyslipidemia. Am J Med.
1009 2016;129(10):1060-6.

1010 81. Martinez-Gomez D, Eisenmann JC, Gomez-Martinez S, Veses A, Romeo J, Veiga OL, et al.
1011 Associations of physical activity and fitness with adipocytokines in adolescents: the AFINOS Study. Nutr
1012 Metab Cardiovasc Dis. 2012;22(3):252-9.

1013 82. Bird SR, Hawley JA. Update on the effects of physical activity on insulin sensitivity in humans. BMJ
1014 Open Sport Exerc Med. 2016;2(1):e000143.

1015 83. Henderson M, Benedetti A, Barnett TA, Mathieu ME, Deladoey J, Gray-Donald K. Influence of
1016 Adiposity, Physical Activity, Fitness, and Screen Time on Insulin Dynamics Over 2 Years in Children. JAMA
1017 Pediatr. 2016;170(3):227-35.

1018 84. Gill JM. Physical activity, cardiorespiratory fitness and insulin resistance: a short update. Curr Opin
1019 Lipidol. 2007;18(1):47-52.

1020 85. Hamer M, O'Donovan G. Cardiorespiratory fitness and metabolic risk factors in obesity. Curr Opin
1021 Lipidol. 2010;21(1):1-7.

1022 86. Nystrom CD, Henriksson P, Martinez-Vizcaino V, Medrano M, Cadenas-Sanchez C, Arias-Palencia NM,
1023 et al. Does Cardiorespiratory Fitness Attenuate the Adverse Effects of Severe/Morbid Obesity on
1024 Cardiometabolic Risk and Insulin Resistance in Children? A Pooled Analysis. Diabetes Care.
1025 2017;40(11):1580-7.

87. Aune D, Norat T, Leitzmann M, Tonstad S, Vatten LJ. Physical activity and the risk of type 2 diabetes: a systematic review and dose-response meta-analysis. Eur J Epidemiol. 2015;30(7):529-42.

88. Rohling M, Strom A, Bonhof G, Puttgen S, Bodis K, Mussig K, et al. Differential Patterns of Impaired Cardiorespiratory Fitness and Cardiac Autonomic Dysfunction in Recently Diagnosed Type 1 and Type 2 Diabetes. Diabetes Care. 2017;40(2):246-52.

89. Juraschek SP, Blaha MJ, Blumenthal RS, Brawner C, Qureshi W, Keteyian SJ, et al. Cardiorespiratory fitness and incident diabetes: the FIT (Henry Ford Exercise Testing) project. Diabetes Care. 2015;38(6):1075-81.

90. Myers J, McAuley P, Lavie CJ, Despres JP, Arena R, Kokkinos P. Physical activity and cardiorespiratory fitness as major markers of cardiovascular risk: their independent and interwoven importance to health status. Prog Cardiovasc Dis. 2015;57(4):306-14.

91. O'Donovan G, Hillsdon M, Ukoumunne OC, Stamatakis E, Hamer M. Objectively measured physical activity, cardiorespiratory fitness and cardiometabolic risk factors in the Health Survey for England. Prev Med. 2013;57(3):201-5.

92. Moore SC, Patel AV, Matthews CE, Berrington de Gonzalez A, Park Y, Katki HA, et al. Leisure time physical activity of moderate to vigorous intensity and mortality: a large pooled cohort analysis. PLoS Med. 2012;9(11):e1001335.

93. Mandsager K, Harb S, Cremer P, Phelan D, Nissen SE, Jaber W. Association of Cardiorespiratory Fitness With Long-term Mortality Among Adults Undergoing Exercise Treadmill Testing. JAMA Netw Open. 2018;1(6):e183605.

94. Chudasama YV, Khunti KK, Zaccardi F, Rowlands AV, Yates T, Gillies CL, et al. Physical activity, multimorbidity, and life expectancy: a UK Biobank longitudinal study. BMC Med. 2019;17(1):108.

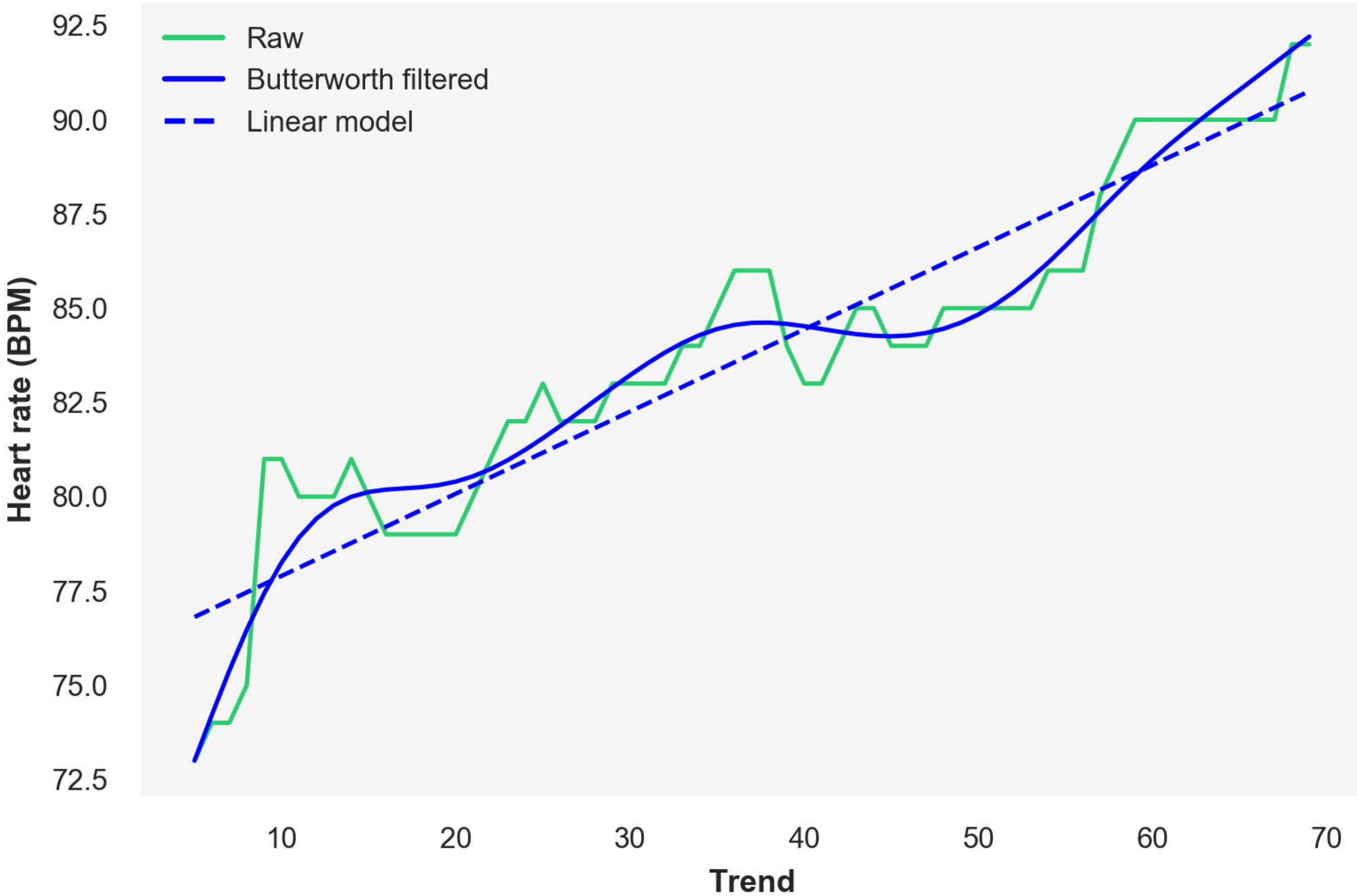
95. Lindgren M, Borjesson M, Ekblom O, Bergstrom G, Lappas G, Rosengren A. Physical activity pattern, cardiorespiratory fitness, and socioeconomic status in the SCAPIS pilot trial - A cross-sectional study. Prev Med Rep. 2016;4:44-9.

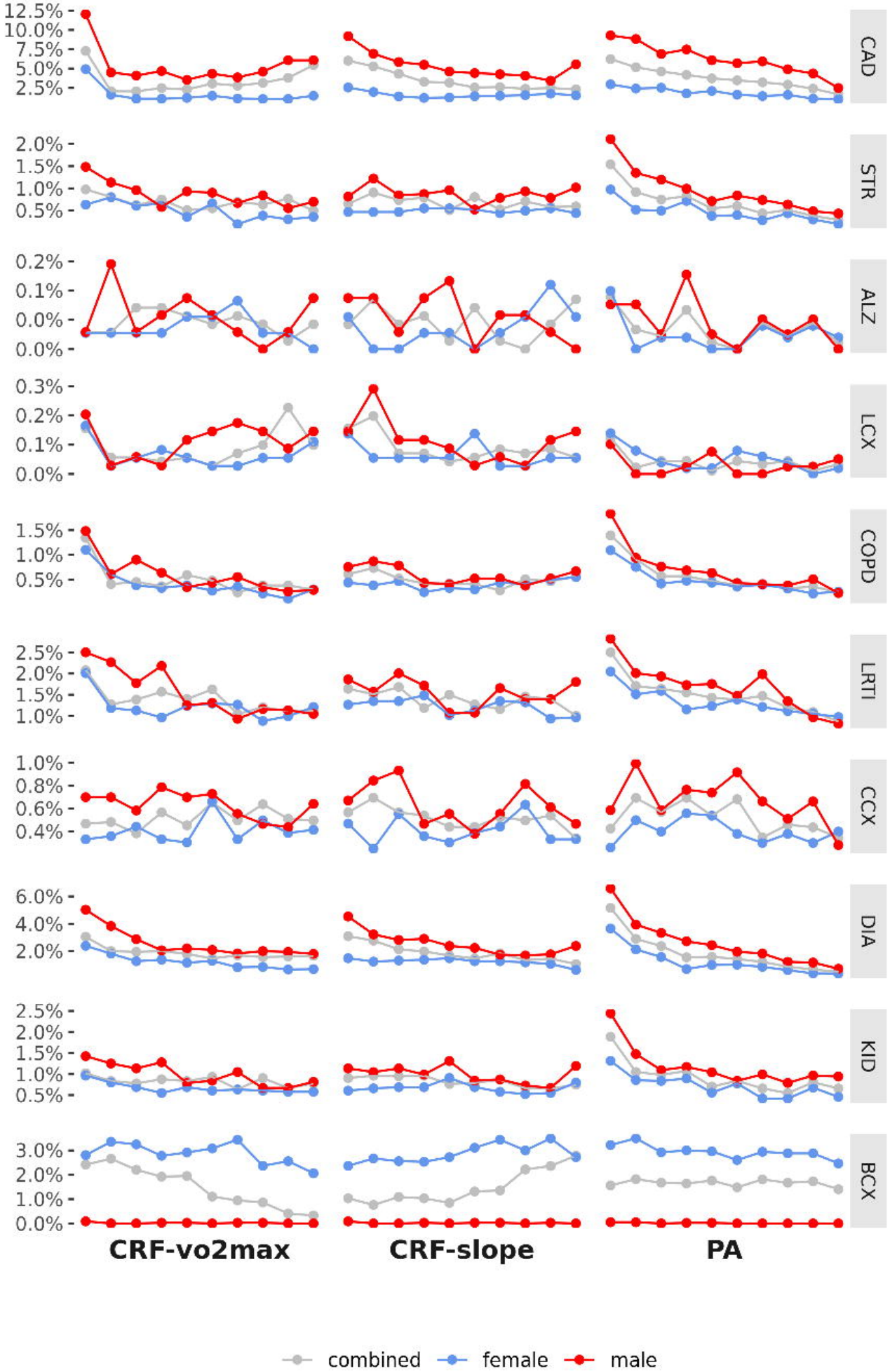
96. Hill JO, Peters JC. Environmental contributions to the obesity epidemic. Science. 1998;280(5368):1371-4.

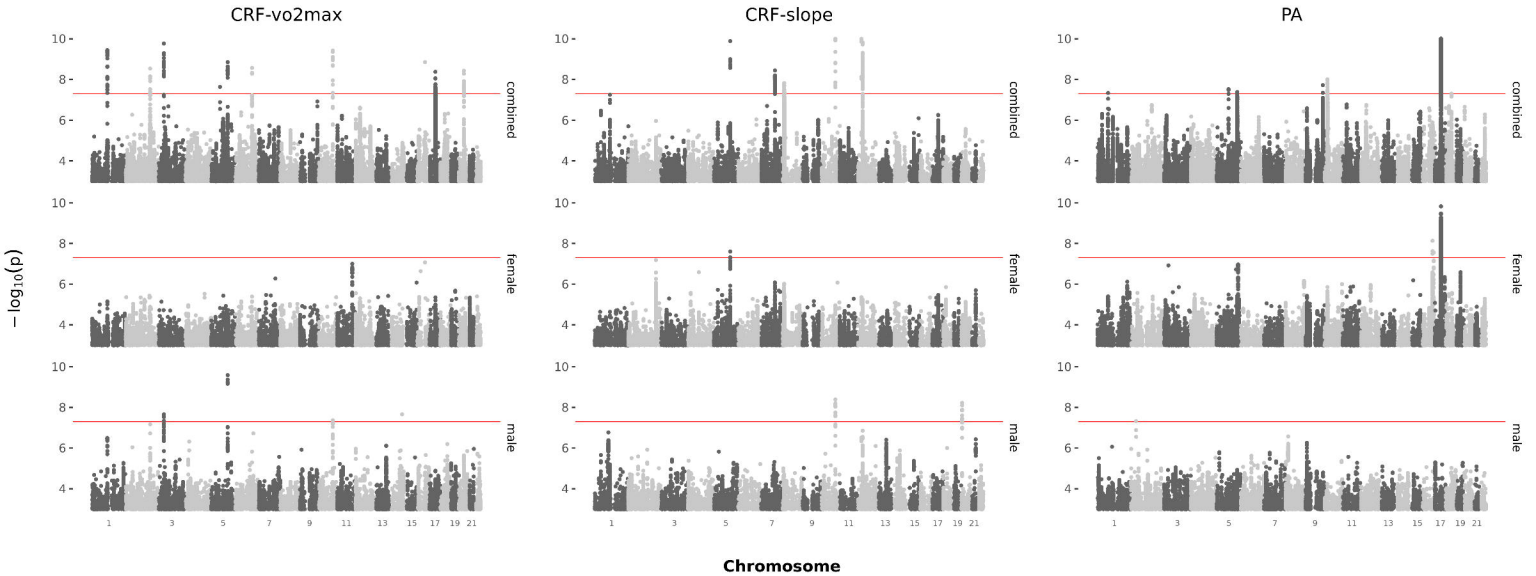
97. French SA, Story M, Jeffery RW. Environmental influences on eating and physical activity. Annu Rev Public Health. 2001;22:309-35.

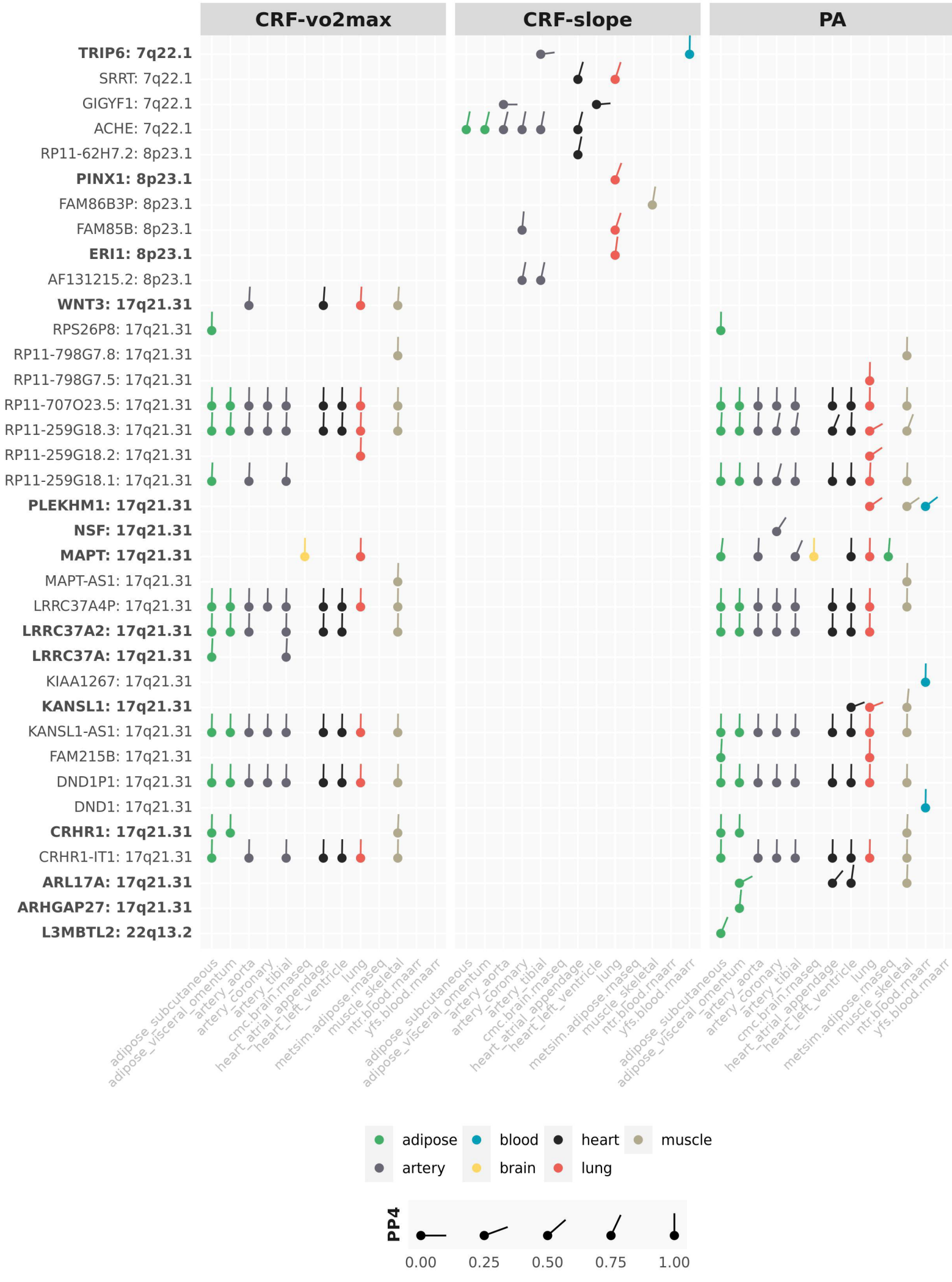
98. Wright N, Wilson L, Smith M, Duncan B, McHugh P. The BROAD study: A randomised controlled trial using a whole food plant-based diet in the community for obesity, ischaemic heart disease or diabetes. Nutr Diabetes. 2017;7(3):e256.

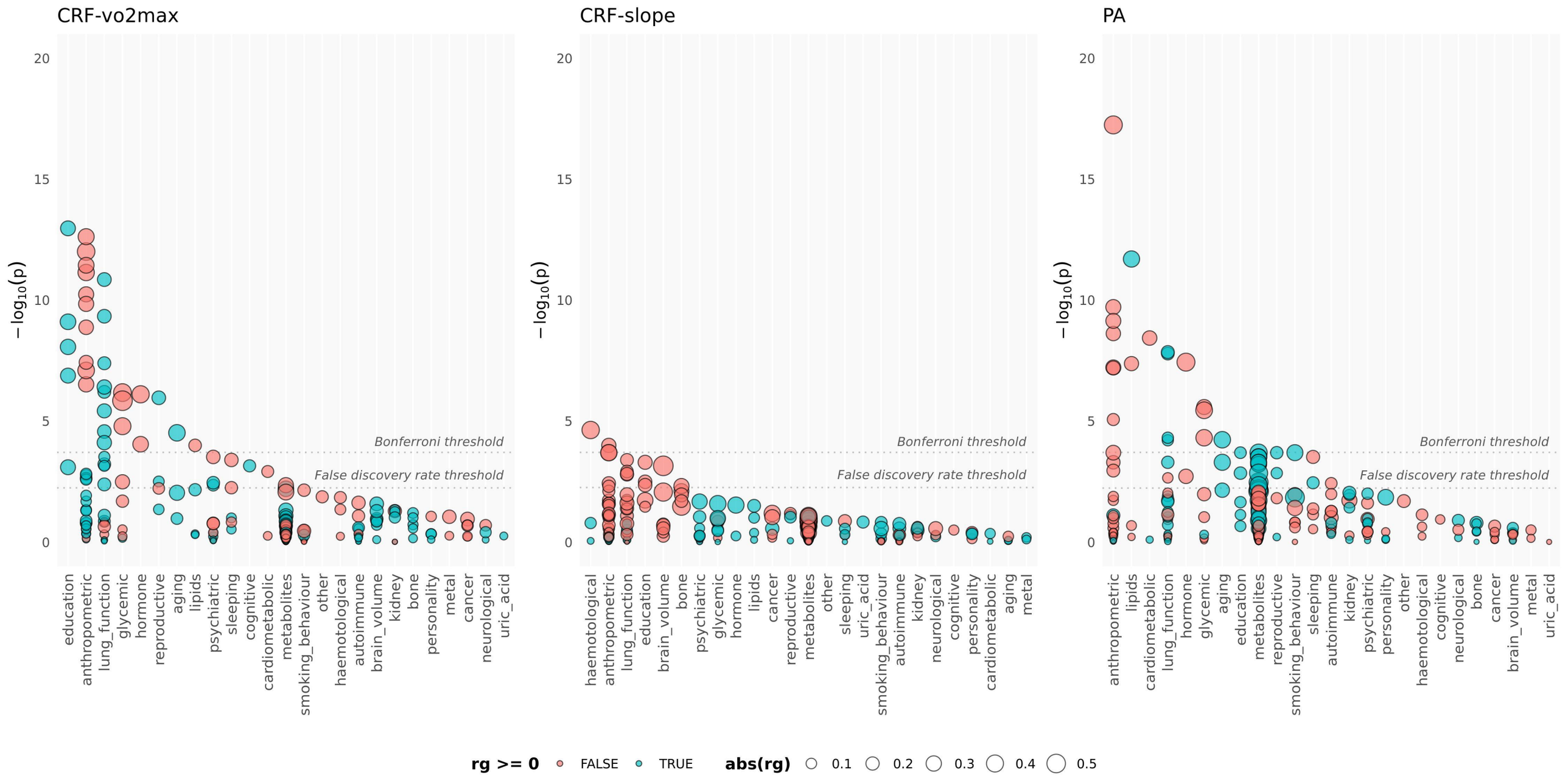
99. Richmond RC, Davey Smith G, Ness AR, den Hoed M, McMahon G, Timpson NJ. Assessing causality in the association between child adiposity and physical activity levels: a Mendelian randomization analysis. PLoS Med. 2014;11(3):e1001618.


Raw
Butterworth filtered
Linear model
Heart rate (BPM)
92.5
90.0
87.5
85.0
82.5
80.0
77.5
75.0
72.5
10
20
30
40
50
60
70
Trend


CAD
STR
ALZ
LCX
COPD
LRTI
CCX
DIA
KID
BCX
CRF-vo2max
CRF-slope
PA
combined
female
male


CRF-vo2max
CRF-slope
PA
combined
female
male
$-\log_{10}(p)$
Chromosome


CRF-vo2max
CRF-slope
PA
TRIP6: 7q22.1
SRRT: 7q22.1
GIGYF1: 7q22.1
ACHE: 7q22.1
RP11-62H7.2: 8p23.1
PINX1: 8p23.1
FAM86B3P: 8p23.1
FAM85B: 8p23.1
ERI1: 8p23.1
AF131215.2: 8p23.1
WNT3: 17q21.31
RPS26P8: 17q21.31
RP11-798G7.8: 17q21.31
RP11-798G7.5: 17q21.31
RP11-707O23.5: 17q21.31
RP11-259G18.3: 17q21.31
RP11-259G18.2: 17q21.31
RP11-259G18.1: 17q21.31
PLEKHM1: 17q21.31
NSF: 17q21.31
MAPT: 17q21.31
MAPT-AS1: 17q21.31
LRRC37A4P: 17q21.31
LRRC37A2: 17q21.31
LRRC37A: 17q21.31
KIAA1267: 17q21.31
KANSL1: 17q21.31
KANSL1-AS1: 17q21.31
FAM215B: 17q21.31
DND1P1: 17q21.31
DND1: 17q21.31
CRHR1: 17q21.31
CRHR1-IT1: 17q21.31
ARL17A: 17q21.31
ARHGAP27: 17q21.31
L3MBTL2: 22q13.2
adipose_subcutaneous
adipose_visceral_omentum
artery_aorta
artery_coronary
artery_tibial
cmc.brain.rnaseq
heart_atrial_appendage
heart_left_ventricle
lung
metsim.adipose.rnaseq
muscle_skeletal
ntr.blood.rnaarr
yfs.blood.rnaarr
adipose
blood
heart
muscle
artery
brain
lung
PP4
0.00
0.25
0.50
0.75
1.00


CRF-vo2max
CRF-slope
PA
$-\log_{10}(p)$
Bonferroni threshold
False discovery rate threshold
education
anthropometric
lung_function
glycemic
hormone
reproductive
aging
lipids
psychiatric
sleeping
cognitive
cardiometabolic
metabolites
smoking_behaviour
other
haemotological
autoimmune
brain_volume
kidney
bone
personality
metal
cancer
neurological
uric_acid
rg >= 0
FALSE
TRUE
abs(rg)
0.1
0.2
0.3
0.4
0.5