Supplementary Methods

**Details on sociodemographic and lifestyle measures**

Age (data field: 21022): Age at recruitment is a derived variable based on date of birth and date of attending an initial Assessment Center. It refers to the age of the participant on the day they attended an initial Assessment Center, truncated to whole year.

Sex (data field: 31): Sex of participant is acquired from central registry at recruitment.

BMI (data field: 21001): BMI value is constructed from height and weight measured during the initial Assessment Center visit. Value is not present if either of these readings were omitted.

Townsend deprivation index (data field: 22189): Neighborhood-level socioeconomic status was measured using the Townsend index of material deprivation. Townsend deprivation index calculated immediately prior to participant joining UK Biobank. Based on the preceding national census output areas. Each participant is assigned a score corresponding to the output area in which their postcode is located.

Education level (data field: 6138): Education level categories were transformed into the following numerical values: “College or University degree” = 20 years; “levels/AS levels or equivalent” = 13 years; “O levels/GCSEs or equivalent” = 10 years; “CSEs or equivalent” = 10 years; “NVQ or HND or HNC or equivalent” = 19 years; “Other professional qualifications, e.g., nursing, teaching” = 15 years; “None of the above” = 7 years; and “Prefer not to answer” = missing.

Current tobacco smoking (data field: 1239): Participants were asked about their current tobacco smoking status. Possible answers were reordered to 0 (No), 1 (Only occasionally), and 2 (Yes, on most or all days). Current tobacco smoking status was treated as a categorical variable. Responses of “prefer not to answer” were excluded.

Alcohol frequency (data field: 1558): Individuals were asked by the touchscreen questionnaire: “About how often do you drink alcohol?” and could respond with, “Never”, “Special occasions only”, “One to three times a month”, “Once or twice a week”, “Three to four times a week” or “Daily or almost daily”. If participants felt that this value varied, they were instructed to give an average over the last year. Responses of “prefer not to answer” were excluded.