**Supplementary figure e-1:** participant flow through the study. The number of individuals at each stage used for analysis is highlighted in the boxes with black outline. The boxes with grey outline indicate the filtering criterion/criteria used at each stage.

Diagram

Description automatically generated

**Supplementary figure e-2:** Nagelkerke’s Pseudo-R2 metric for all 32 MHC PRS without controlling for principal components.

Chart, bar chart

Description automatically generated

**Supplementary figure e-3:** Nagelkerke’s Pseudo-R2 metric for all 32 MHC PRS controlling for ten principal components.

Chart, bar chart

Description automatically generated

**Supplementary figure e-4:** density plot depicting age at MS diagnostic code report. Note that these codes refer to the first reported diagnosis, not the first symptom onset, and are at best a proxy for disease onset.

A group of people in different colors

Description automatically generated

**Supplementary figure e-5:** no association between PRS and claiming of disability benefits among people with MS.

A picture containing schematic

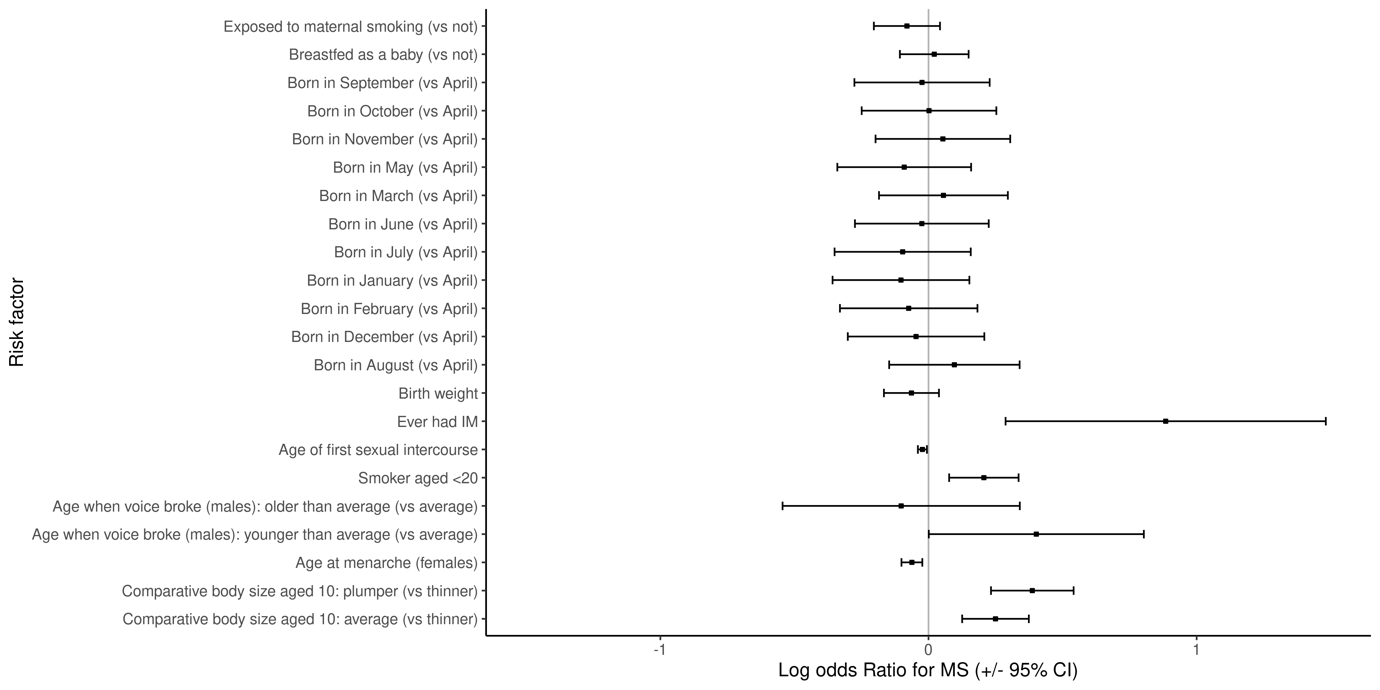
Description automatically generated

**Supplementary figure e-6:** multiplicative interaction terms and 95% confidence intervals for interaction between environmental exposures and MS genetic risk (MHC and non-MHC scores).

Chart, box and whisker chart

Description automatically generated

**Supplementary figure e-7:** sensitivity analysis – association between exposure variables and MS in UK Biobank having excluded MS cases defined only by a single source of report. ORs and CIs are from the output of a multivariable logistic regression with the following covariates: age, sex, ethnicity, birth latitude, current deprivation status, and the exposure in question. For menarche (females only) and voice-breaking (males-only), sex was not included as a covariate.



**Supplementary figure e-8:** performance characteristics of the MHC and non-MHC polygenic risk scores in a sensitivity analysis restricting the MS case definition to those individuals who had at least two sources of MS report (i.e. two of death register, self-report, HES code, or primary care).

(a)Training set Nagelkerke pseudo-R2 values,

(b) validation set PRS distributions,

© validation set calibration,

(d) validation set discrimination characteristics are shown (in order).

(a)

Chart, bar chart

Description automatically generated

(b)

Chart, bar chart, histogram

Description automatically generated

(c)

Chart

Description automatically generated

(d)

Diagram

Description automatically generated

**Supplementary figure e-9:** estimates for the Attributable Proportion due to interaction (AP) and 95% confidence intervals for each PRS x environmental exposure interaction tested.

