Supplementary Appendix 1: Case definitions and the International Classification of Diseases diagnosis codes used to ascertain the study outcomes

| Outcome | Case definition | ICD codes |
|---|---|---|
| Mental disorders | Fulfilling the case definition of either mood and anxiety, psychotic or substance use disorders as described below | |
| Mood and anxiety disorders | At least one hospitalisation or two physician visits within three years | ICD-8: 296, 300, 307<br>ICD-9-CM: 296, 311, 309, 300<br>ICD-10-CA: F30-F34, F38, F40, F41, F41.1, F41.2, F41.3, F41.8, F41.9, F42, F43, F53 |
| Psychotic disorders | At least one hospitalisation or one physician visit | ICD-8: 295, 297, 298, 299<br>ICD-9-CM: 295, 297, 298<br>ICD-10-CA: F11.5, F12.5, F13.5, F14.5, F15.5, F16.5, F17.5, F18.5, F19.5, F20-F25, F28, F29 |
| Substance use disorders | At least one hospitalisation or one physician visit | ICD-8: 291, 294, 303, 304<br>ICD-9-CM: 291, 292, 303-305<br>ICD-10-CA: F10-19, F55, Z50.2, Z50.3 |
| Ischaemic heart disease | At least one hospitalisation or two physician visits within one year | ICD-8: 410-414<br>ICD-9-CM: 410-414<br>ICD-10-CA: I20-I25 |

ICD: International Classification of Diseases, ICD-8: the eighth revision of ICD, ICD-9-CM: the ninth revision of ICD with clinical modifications, ICD-10-CA: the tenth revision of ICD with Canadian adaptations.

Supplementary Appendix 2: The number and percentage of offspring with and without any candidate fathers in the pre-1980 cohort across missingness levels

| Simulated missingness level (%) | 10 | 20 | 30 | 40 | 50 |
|---|---|---|---|---|---|
| With at least 1 candidate father, N (%) | 2638 (78.9) | 5267 (78.8) | 7773 (77.5) | 10,420 (77.9) | 13,070 (78.2) |
| Without any candidate fathers, N (%) | 704 (21.1) | 1418 (21.2) | 2254 (22.5) | 2950 (22.1) | 3643 (21.8) |

Supplementary Appendix 3: The percentage of candidate father sets that included the true father in the pre-1980 cohort, before and after restricting sets to three fathers

| Simulated missingness level (%) | 10 | 20 | 30 | 40 | 50 |
|---|---|---|---|---|---|
| Candidate sets limited to 3 fathers, % | 26.5 | 27.9 | 28.2 | 28.3 | 28.2 |
| Unrestricted candidate sets, % | 64.8 | 65.8 | 65.7 | 66.1 | 66.0 |

Supplementary Appendix 4: Hazard ratios (HRs) and 95% confidence intervals (CIs) for the association of paternal history of mental disorders or ischaemic heart disease (IHD) with their risks in offspring using true fathers, candidate fathers, indicator category and complete-case approaches in the pre-1980 cohort

| Missingness level (%) | HR (95% CI) | | | |
|---|---|---|---|---|
| | True fathers N = 33,426 | Candidate father approach* | Indicator category (N = 33,426) | Complete-case** |
| **Mental disorders** | | | | |
| 10 | 1.27 (1.23-1.31) | 1.24 (1.20-1.29) N = 32,722 | 1.26 (1.22-1.31) | 1.26 (1.22-1.31) N = 30,084 |
| 20 | | 1.21 (1.17-1.26) N = 32,008 | 1.26 (1.21-1.30) | 1.26 (1.21-1.30) N = 26,741 |
| 30 | | 1.20 (1.15-1.25) N = 31.172 | 1.27 (1.22-1.32) | 1.27 (1.22-1.32) N = 23,399 |
| 40 | | 1.19 (1.14-1.23) N = 30,476 | 1.28 (1.22-1.33) | 1.28 (1.23-1.34) N = 20,056 |
| 50 | | 1.17 (1.13-1.21) N = 29,783 | 1.28 (1.22-1.34) | 1.28 (1.22-1.33) N = 16,713 |
| **IHD** | | | | |
| 10 | 1.94 (1.35-2.79) | 1.93 (1.33-2.80) N = 32,722 | 1.99 (1.37-2.90) | 2.01 (1.38-2.93) N = 30,084 |
| 20 | | 1.96 (1.31-2.94) N = 32,008 | 2.04 (1.38-3.00) | 2.08 (1.41-3.07) N = 26,741 |
| 30 | | 1.87 (1.25-2.80) N = 31.172 | 2.11 (1.40-3.18) | 2.15 (1.43-3.25) N = 23,399 |
| 40 | | 1.71 (0.95-3.08) N = 30,476 | 1.99 (1.26-3.14) | 2.02 (1.28-3.19) N = 20,056 |
| 50 | | 1.65 (0.85-3.20) N = 29,783 | 2.26 (1.39-3.67) | 2.32 (1.43-3.77) N = 16,713 |

*Excluded those without any candidate fathers, which varies at each missingness level.
**Excluded those with artificially excluded fathers (simulated missingness).

Supplementary Appendix 5: Hazard ratios (HRs) and 95% confidence intervals (CIs) for the association of paternal history with offspring mental disorders and ischaemic heart disease (IHD) in the pre-1980 validation cohort, stratified by whether the selected candidate father matched the true father (20% simulated missingness)

| Outcome | HR (95% CI)* | | | |
|---|---|---|---|---|
| | True fathers N = 33,426 | All candidate fathers | Candidate father was the true father | Candidate father was not the true father |
| **Mental disorders** | 1.27 (1.23–1.32) | 1.22 (1.18–1.26) | 1.26 (1.06–1.51) | 1.01 (0.92–1.11) |
| **IHD** | 1.94 (1.35–2.79) | 1.83 (1.25–2.69) | NA** | NA** |

*Birth decade was excluded from the models because there was no variation in its distribution among offspring whose selected candidate father matched or did not match the true father.
**Analyses could not be run because of the small number of events.

Supplementary Appendix 6: Concordance of paternal history classification with the true father among offspring with simulated missing paternal linkage (20% simulated missingness) in the pre-1980 validation cohort for mental disorders and ischaemic heart disease (IHD)

| Paternal History | Candidate Father Iteration | True Yes/ Candidate Yes | True Yes/ Candidate No | True No/ Candidate Yes | True No/ Candidate No |
|---|---|---|---|---|---|
| Mental disorders | 1 | 860 (16.3%) | 850 (16.1%) | 913 (17.3%) | 2644 (50.2%) |
| | 2 | 603 (12.7%) | 928 (19.5%) | 1022 (21.5%) | 2210 (46.4%) |
| | 3 | 487 (11.6%) | 873 (20.8%) | 948 (22.6%) | 1895 (45.1%) |
| IHD | 1 | 82 (1.6%) | 227 (4.3%) | 214 (4.1%) | 4744 (90.1%) |
| | 2 | 37 (0.8%) | 231 (4.9%) | 222 (4.7%) | 4273 (89.7%) |
| | 3 | 18 (0.4%) | 216 (5.1%) | 219 (5.2%) | 3750 (89.2%) |

Note: reported as frequencies (percentages).

Supplementary Appendix 7: Characteristics of offspring with simulated missing paternal linkage according to whether an eligible candidate father was identified in the pre-1980 validation cohort (20% simulated missingness)

| Characteristic | Candidate father identified (5267, 16.5%) | No eligible candidate father (26741, 83.5%) |
|---|---|---|
| Male sex | 2760 (52.4%) | 13886 (51.9%) |
| Urban residence | 2754 (52.3%) | 14914 (55.8%) |
| Income quintile | | |
| 1 (lowest) | 1016 (19.3%) | 4925 (18.4%) |
| 2 | 1349 (25.6%) | 6863 (25.7%) |
| 3 | 1193 (22.7%) | 6150 (23.0%) |
| 4 | 922 (17.5%) | 4890 (18.3%) |
| 5 (highest) | 595 (11.3%) | 3083 (11.5%) |
| Missing | 192 (3.7%) | 830 (3.1%) |
| Mental disorder diagnosis | 2440 (46.3%) | 12378 (46.3%) |
| IHD diagnosis | 33 (0.6%) | 252 (0.9%) |
| Number of comorbidities* | | |
| None | 3472 (65.9%) | 17980 (67.2%) |
| 1 | 1616 (30.7%) | 7807 (29.2%) |
| 2+ | 179 (3.4%) | 954 (3.6%) |
| Maternal mental disorder history | 2505 (47.6%) | 12481 (46.7%) |
| Maternal IHD history | 83 (1.6%) | 422 (1.6%) |
| Maternal age | | |
| <20 | 422 (8.0%) | 1922 (7.2%) |
| 20-24 | 1404 (26.7%) | 7338 (27.4%) |
| 25-29 | 2035 (38.6%) | 10013 (37.4%) |
| 30+ | 1406 (26.7%) | 7468 (27.9%) |

Abbreviations: IHD: ischaemic heart disease.
Note: reported as frequencies (percentages).
*Based on the 17 chronic conditions in the Charlson Comorbidity Index.

Supplementary Appendix 8: Hazard ratios (HRs) and 95% confidence intervals (CIs) for the association between paternal history of mental disorders and ischaemic heart disease (IHD) and offspring outcomes in the full cohort, stratified by ICD coding period

| Outcome/ICD coding period | HR* (95% CI) | | |
|---|---|---|---|
| | Candidate father | Indicator category | Complete-case |
| Mental disorders | 1.21 (1.19-1.24) | 1.26 (1.23-1.28) | 1.37 (1.34-1.40) |
| Index date <April 1, 2004 | 1.22 (1.19-1.25) | 1.25 (1.21-1.28) | 1.39 (1.35-1.44) |
| Index date >=April 1, 2004 | 1.20 (1.15-1.25) | 1.29 (1.24-1.33) | 1.35 (1.31-1.40) |
| IHD | 1.71 (1.23-2.39) | 1.83 (1.33-2.51) | 1.35 (1.09-1.68) |
| Index date <April 1, 2004 | 1.76 (1.25-2.49) | 1.82 (1.31-2.53) | 1.31 (1.02-1.67) |
| Index date >=April 1, 2004 | 1.42 (0.49-4.16) | 2.06 (0.74-5.76) | 1.52 (0.94-2.47) |

*Birth decade was excluded from the models because there was no variation in its distribution among offspring with index date before or after April 1, 2004.