**Appendix B: Data dictionary for candidate predictors**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **No.** | **Variable name** | **Description** | **Format** | **Response options** | **Note** |
|  | bmi | Women’s BMI (kg/m2) | Continuous | 99 = Not stated/missing | Weights were collected at different time points in NPDC (record in the first trimester of pregnancy), VPDC (self-reported at conception) and QPDC (self-reported four to six weeks prior to or at conception). |
|  | cob\_region | Women’s country of birth grouped into broad regions | Categorical | 1 = Australia (includes External Territories)  2 = Oceania and antarctica (excludes Australia)  3 = Europe  4 = Africa  5 = Asia  6 = America  9 = Not stated/missing | Women’s country of birth was categorized based on the Standard Australian Classification of Countries (SACC) used by the Australian Bureau of Statistics (ABS) (1). |
|  | gest\_diab | Gestational diabetes during current pregnancy | Dichotomous | 1 = Yes  0 = No/ not stated/missing | As collected and defined in the NPDC and VPDC.  O244 (ICD-10-AM code) reported in pregnancy complications collected in the QPDC. |
|  | gest\_hyper | Gestational hypertension during current pregnancy | Dichotomous | 1 = Yes  0 = No/ not stated/missing | As collected and defined in the NPDC and VPDC.  O13 (ICD-10-AM code) reported in pregnancy complications in the QPDC. |
|  | gest\_weeks | Weeks of gestation at birth | Categorical | 37 = 37+0 - 37 +6  38 = 38+0 - 38+6  39 = 39+0 - 39+6  40 = 40+0 - 40+6  41 = 41+0 - 41+6  99 = Not stated/missing | As collected and defined in the NPDC, VPDC and QPDC. |
|  | gest\_weeks\_first\_visit | Weeks of gestation at first antenatal visit | Continuous | 99 = Not stated/missing | As collected and defined in the NPDC, VPDC and QPDC. |
|  | gravidity | Number of previous pregnancies | Continuous | 99 = Not stated/missing | As collected and defined in the NPDC, VPDC and QPDC. |
|  | mother\_age | Women’ age at time of giving birth | Continuous | 99 = Not stated/missing | As collected and defined in the NPDC, VPDC and QPDC. |
|  | num\_ant\_visit | Number of total antenatal visits | Continuous | 99 = Not stated/missing | As collected and defined in the NPDC, VPDC and QPDC. |
|  | parity | Whether women gave birth before current pregnancy (live births or stillbirths of at least 20 weeks of gestation and/or at least 400 grams in weight) | Categorical | 0 = No (nulliparous)  1 = Yes (parous with prior vaginal birth)  9 = Not stated/missing | As collected and defined in the NPDC, VPDC and QPDC. |
|  | pre\_diab | Pre-existing diabetes (any type) | Dichotomous | 1 = Yes  0 = No/ not stated/missing | As collected and defined in the NPDC and VPDC.  E10-E14 or O24 (ICD-10-AM code) reported in current medical conditions in the QPDC. |
|  | pre\_hyper | Pre-existing hypertension (any type) | Dichotomous | 1 = Yes  0 = No/ not stated/missing | As collected and defined in the NPDC and VPDC.  I11-I15 or O10 or O11 (ICD-10-AM code) reported in current medical conditions in the QPDC. |
|  | preeclampsia | Preeclampsia during current pregnancy | Dichotomous | 1 = Yes  0 = No/ not stated/missing | As collected and defined in the NPDC and VPDC.  O11, O14 or O15 (ICD-10-AM code) reported in current medical conditions in the QPDC. |
|  | private\_hospital | Giving birth in private or public hospitals | Categorical | 1 = Yes  0 = No  9 = Not stated/missing | As collected and defined in the NPDC, VPDC and QPDC. |
|  | rurality | Women’s residential area’s rurality group | Ordinal | 1 = Major cities of Australia  2 = Inner regional Australia  3 = Outer regional Australia  4 = Remote Australia  5 = Very remote Australia  9 = Not stated/missing/migratory-offshore-shipping/no usual address | Rurality was categorized based on the postcode of usual residence using the Accessibility/Remoteness Index of Australia (ARIA+) used by the Australian Bureau of Statistics (ABS) (2).  2011 release was used by the NPDC. 2016 release was used for VPDC and QPDC. |
|  | seifa | Women’s residential area of socioeconomic disadvantage | Ordinal | 1 = Quintile 1 (the most disadvantaged)  2 = Quintile 2  3 = Quintile 3  4 = Quintile 4  5 = Quintile 5 (the least disadvantaged)  9 = Not stated/missing | Socioeconomic status was categorized based on the postcode of usual residence using the Index of Relative Socio-economic Disadvantage (IRSD) in the Socio-Economic Indexes for Areas (SEIFA) - a classification system developed by the Australian Bureau of Statistics (ABS) that ranks geographic areas across Australia according to relative socio-economic advantage and disadvantage (3).  2011 release was used by the NPDC. 2016 release was used for VPDC and QPDC. |
|  | smoke\_a20 | Smoke after 20 weeks of gestation | Categorical | 1 = Yes  0 = No  9 = Not stated/missing | As collected and defined in the NPDC and VPDC. |
|  | smoke\_b20 | Smoke before 20 weeks of gestation | Categorical | 1 = Yes  0 = No  9 = Not stated/missing | As collected and defined in the NPDC and VPDC. |

Abbreviations: BMI=Body Mass Index; NPDC=New South Wales Perinatal Data Collection; VPDC=Victoria Perinatal Data Collection; QPDC=Queensland Perinatal Data Collection.

**References**

1. Australian Bureau of Statistics. Standard Australian Classification of Countries (SACC) Canberra: ABS; 2016 [2024 August 14]. Available from: <https://www.abs.gov.au/statistics/classifications/standard-australian-classification-countries-sacc/latest-release>.

2. Australian Bureau of Statistics. Remoteness Areas Canberra: ABS; Jul2021-Jun2026 [2024 August 14]. Available from: <https://www.abs.gov.au/statistics/standards/australian-statistical-geography-standard-asgs-edition-3/jul2021-jun2026/remoteness-structure/remoteness-areas>.

3. Australian Bureau of Statistics. Socio-Economic Indexes for Areas (SEIFA), Australia: ABS; 2021 [Available from: <https://www.abs.gov.au/statistics/people/people-and-communities/socio-economic-indexes-areas-seifa-australia/2021>.