

1 **Original Research**

2 **Maternal cardiovascular and haematological complications alter the risk associations**
3 **between environmental exposure and adverse pregnancy outcomes**

4
5 **Jason Sun**^{1,2,3,4,5*}, **Haiyang Tang**¹, **Huan Zhao**^{6,1}, **Qingyi Xiang**¹, **Yijia Tian**¹, **Kim Robin van Daalen**^{5,7,8},
6 **Kun Tang**⁹, **Evelyn Xiu-Ling Loo**^{2,10,11}, **Lynette P Shek**^{2,3}, **Alexander T Archibald**^{1,12}, **Wei Xu**¹³, **Yuming**
7 **Guo**^{14*}, and **Xiaoxia Bai**^{1,15,16,17*}, on behalf of the Zhejiang Environmental and Birth Health Research Alliance
8 (ZEBRA) collaborative group

9

- 10 ¹ Department of Obstetrics, Women's Hospital, Zhejiang University School of Medicine, Hangzhou 310006, Zhejiang Province, PR China
11 ² Department of Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117609, Republic of Singapore
12 ³ Centre for Sustainable Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117609, Republic of
13 Singapore
14 ⁴ Yusuf Hamied Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, UK
15 ⁵ British Heart Foundation Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge,
16 Cambridge CB1 8RN, UK
17 ⁶ Department of Obstetrics and Gynecology, The Fourth Affiliated Hospital, Zhejiang University School of Medicine, Yiwu, Zhejiang 322000,
18 PR China
19 ⁷ Heart and Lung Research Institute, University of Cambridge, Cambridge CB2 0BD, UK
20 ⁸ Barcelona Supercomputing Centre, Department of Earth Sciences, Barcelona 08034, Spain
21 ⁹ Vanke School of Public Health, Tsinghua University, Beijing 100084, China
22 ¹⁰ Singapore Institute for Clinical Sciences (SICS), Agency for Science, Technology and Research (A*STAR), 30 Medical Drive, Singapore
23 117609, Republic of Singapore
24 ¹¹ Human Potential Translational Research Programme, Yong Loo Lin School of Medicine, National University of Singapore, Singapore,
25 Republic of Singapore
26 ¹² National Centre for Atmospheric Science, Cambridge CB2 1EW, UK
27 ¹³ Maternal and Child Health Division, Health Commission of Zhejiang Province, Hangzhou 310006, Zhejiang Province, PR China
28 ¹⁴ School of Public Health and Preventive Medicine, Monash University, Melbourne, VIC 3004, Australia
29 ¹⁵ Traditional Chinese Medicine for Reproductive Health Key Laboratory of Zhejiang Province, Hangzhou 310006, Zhejiang Province, PR
30 China
31 ¹⁶ Zhejiang Provincial Clinical Research Centre for Obstetrics and Gynaecology, Hangzhou 310006, Zhejiang Province, PR China
32 ¹⁷ Key Laboratory of Women's Reproductive Health, Hangzhou 310006, Zhejiang Province, PR China

33

34 *Correspondence to:

35 yuming.guo@monash.edu (YG), baixiaoxia@zju.edu.cn (XB) and csuen27@gmail.com (JS).

36

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

37 **Given China’s recent introduction of the “three-child policy” in response to population ageing¹,**
38 **safeguarding perinatal health has become an urgent priority². Previous epidemiological research seldom**
39 **explored the risk factors of maternal cardiovascular and haematological diseases, or its impact on adverse**
40 **pregnancy outcomes (APO). To fill the literature gap, here we conducted systematic epidemiological**
41 **analyses on 121,090 pregnant women and their neonates from the ZEBRA Chinese prospective maternity**
42 **cohort. We find that incremental exposure in PM_{2.5}, O₃, and green space modify the risks of APO,**
43 **including congenital heart disease, by 11.2%, 7.8%, and –5.5%, respectively. Maternal cardiovascular**
44 **and haematological complications during pregnancy significantly aggravate the risk of APO by 66.2%,**
45 **and also modify the environment-APO risk associations by amplifying the hazards of air pollution and**
46 **weakening the protective effect of greenness accessibility. Our research findings support the Sustainable**
47 **Development Goals (e.g. SDG3)^{3,4} by providing first-hand epidemiological evidence and clinical guidance**
48 **for protecting maternal and neonatal health.**

49 Studies have assessed the effects of environmental exposures, including ambient air pollution, heatwave,
50 and green space, on the risk of cardiovascular morbidity and mortality⁵⁻¹⁰. However, research that focused on
51 the pregnant woman and the neonate, two groups of vulnerable population, is still limited. Literature has also
52 reported strong associations between maternal exposure to various environmental factors and multiple types of
53 APO¹¹⁻¹⁶, but other pregnancy abnormalities (e.g. neonatal congenital heart disease, CHD) other than preterm
54 birth (PTB), term low birth weight (LBW), and stillbirth have seldom been analysed. It is noteworthy that
55 women with pregnancy-induced cardiovascular diseases (CVDs) have an increased risk of developing
56 hypertension in later years^{17,18}. Research has shown that women with a history of pre-eclampsia have an
57 approximately four-fold higher incidence of stroke in later years¹⁷, as pre-eclampsia can cause maternal vascular
58 remodelling¹⁹. For the neonates diagnosed with CHD, childhood mortality rates are rather high²⁰; even with

59 medical treatment, adults and adolescents often have lower exercise capability²¹, and a portion of women are
60 even advised against carrying to pregnancy to term in fear of cardiac complications and even sudden death of
61 the child²². Although highly sensitive screening tools are now being widely used for timely prenatal detection
62 of CHD²³, identifying risk factors to minimise the occurrence risk of CHD is always the ultimate pursuit.

63 Therefore, population-based epidemiological research on cardiovascular-related outcomes during the
64 perinatal period holds significant public health implications. By systematic analyses of the medical records of
65 pregnant women enrolled in the ZEBRA Chinese maternity cohort²⁴, our current study aims to explore: 1) the
66 risk associations between environmental exposures and pregnancy-induced CVDs and APO; 2) the mediating
67 role of maternal cardiovascular symptoms (both primary and pregnancy-induced) in the environment-APO risk
68 association; and 3) the early-stage predictability of APO integrating maternal diagnosis of CVDs. Besides
69 contributing to filling the literature gap, the risk prediction models also provide the basis for further clinical
70 validation and application. Protecting the mother's health is not only an important element of gender equality,
71 but also a reflection of civilisation. Ensuring the health of newborns is crucial as it directly impacts the lifelong
72 well-being and human potential of the next generation, serving as a vital factor in advancing society.

73

74 **RESULTS**

75 The ZEBRA (Zhejiang Environmental and Birth Health Research Alliance) maternity cohort recruited 137,392
76 Chinese pregnant women nationwide during 2013–2022, among which 121,090 pregnant women were included
77 for epidemiological analyses (Fig. 1). There were 25,544 (21.1%) pregnant women diagnosed with a variety of
78 cardiovascular and haematological complications, including 814 (0.7%) cases of heart disease, 3,760 (3.1%)
79 cases of vascular disease, and 21,935 (18.1%) cases of haematological disease. Pregnancy-induced
80 cardiovascular complications were found in 1,414 (1.2%) cases (Extended Data Fig. 1). After excluding 802

81 stillbirths, there were 9,501 cases of APOs among the remaining 124,025 live births (3,619 pairs of twins and
82 59 sets of triplets), including 6,604 (5.3%) PTB, 4,849 (3.9%) term LBW, 840 (0.7%) CHD, and 3,334 (2.7%)
83 blood disorders (Extended Data Fig. 2). Detailed cohort profile statistics are listed in Extended Data Table 1,
84 and medical history in Extended Data Table 2.

85

86 Environmental risks on perinatal CVDs

87 To investigate the risk associations between environmental exposure and pregnancy-induced cardiovascular
88 complications, we defined the eighteen month period starting from one year before conception till the end of
89 the second trimester as the exposure window²⁵. This yielded individual-level maternal exposure to ambient
90 PM_{2.5}, O₃, and greenness as 40.3±7.4 µg/m³, 47.1±4.0 ppb, and 0.20±0.07 quantified in the enhanced vegetation
91 index (EVI), respectively. Having adjusted for sociodemographic, physiological and behavioural characteristics,
92 medical and disease history, and other environmental exposures (see details in Extended Data Table 3), the risk
93 of pregnancy-induced cardiovascular diseases would increase by 7.3% (hazard ratio, HR=1.073, 95% CI:
94 1.064–1.082, $p=7.83\times 10^{-23}$) with each 10-µg/m³ increase in PM_{2.5} exposure, increase by 2.7% (2.2–3.3%,
95 $p=1.78\times 10^{-12}$) with each 10-ppb incremental O₃ exposure, and reduce by 3.6% (1.8–5.2%, $p=1.47\times 10^{-4}$) with
96 every 0.1-EVI additional green space exposure. Significant risk associations were observed for two specific
97 subsets of cardiovascular diseases, that is, the pregnancy-induced hypertension and preeclampsia (Fig. 2).

98 As for the neonatal cardiovascular diseases, it was found that for every 10-µg/m³ exposure increase in
99 PM_{2.5}, 10-ppb O₃, and 0.1-EVI greenness, there would be 1.1% (0.5–1.8%, $p=1.63\times 10^{-3}$), 1.6% (1.2–2.0%,
100 $p=1.79\times 10^{-10}$), and –1.8% (HR=0.982, 95% CI: 0.970–0.995, $p=4.79\times 10^{-3}$) change in the incidence risks,
101 respectively (Fig. 2). The neonatal cardiac anomalies had the most prominent associations with maternal PM_{2.5}
102 and O₃ exposures. The protective effect of maternal greenness exposure was not significant on the neonatal

103 incidence of CHD Environmental exposures also demonstrated additional risks on other adverse perinatal
104 outcomes such as stillbirth, PTB, term LBW, and respiratory disease, where the quantitative relationships
105 showed monotonically increasing exposure-response tendencies (Extended Data Fig. 3).

106 Synergistic effect between PM_{2.5} and O₃ exposure was observed as 4.2% (effect modification, EM=1.042,
107 95% CI: 1.022–1.062, $p=3.88\times 10^{-5}$) for pregnancy-induced cardiovascular complications, indicating for each
108 10-ppb increase in maternal O₃ exposure, the risk strength between PM_{2.5} exposure (by 10- $\mu\text{g}/\text{m}^3$ increment)
109 and cardiovascular incidence would rise by 4.2%, and *vice versa*. The synergistic hazard was more significant
110 in the incidence risk of APOs (EM=1.054, 95% CI: 1.046–1.061, $p=2.33\times 10^{-23}$). On the contrary, green space
111 accessibility exhibited antagonistic effects with the two major ambient air pollutants, as -9.7% (EM=0.903, 95%
112 CI: 0.883–0.923, $p=2.41\times 10^{-13}$) on the risk association between PM_{2.5} and pregnancy-induced cardiovascular
113 diseases, implying that with each 0.1-EVI incremental greenness exposure, the PM_{2.5}-cardiovascular risk
114 strength scaled in 10- $\mu\text{g}/\text{m}^3$ increase of exposure would be compromised by 9.7%. The antagonism of greenness
115 exposure on O₃ was relatively weaker at -4.4% (EM=0.956, 95% CI: 0.938–0.973, $p=2.56\times 10^{-6}$). The protective
116 effect of greenness against the occurrence risk of APO was 8.6% (7.7–9.4%, $p=2.08\times 10^{-5}$) and 4.7% (4.1–5.4%,
117 $p=2.15\times 10^{-20}$) for PM_{2.5} and O₃ exposure, respectively. The foregoing analyses revealed unneglectable risks
118 associated with ambient air pollution on the well-being of expectant mothers and neonates, particularly in terms
119 of cardiovascular health. Conversely, the greenness accessibility not only contributes to the reduction of
120 perinatal anomalies but also offers a partial amelioration of the adverse impacts from PM_{2.5} and O₃ exposure.
121 From a policy perspective, our findings underscore the importance of synergistic PM_{2.5}-O₃ control and urban
122 greening for public health benefits.

123

124 **Impacts of maternal CVDs on adverse pregnancy outcomes**

125 Maternal primary and pregnancy-induced cardiovascular diseases are significant risk factors for APOs, with an
126 overall HR=1.693 (95% CI: 1.615–1.775, $p=1.70\times 10^{-26}$), indicating pregnant individuals with composite CVDs
127 suffer a 69.3% higher risk to develop adverse gestational outcomes compared to those with normal
128 cardiovascular function. More specifically, pregnancy-induced cardiovascular diseases result in 28.4% higher
129 risks of APOs (HR=1.284, 95% CI: 1.025–1.609, $p=1.77\times 10^{-2}$), whereas primary cardiovascular diseases can
130 lead to a more prominent risk as 37.0% (HR=1.370, 95% CI: 1.304–1.439, $p=6.81\times 10^{-17}$). Fig. 3 presents the
131 association strengths between specific subtypes of maternal cardiovascular disease during pregnancy and APO
132 subsets, revealing in general that pregnancy-induced maternal cardiovascular diseases, particularly hypertension
133 and preeclampsia, exhibit stronger risk associations with various neonatal anomalies compared to primary
134 cardiovascular diseases.

135 Maternal heart diseases could impact neonatal cardiovascular diseases (Fig. 3). Notably, an observable
136 mother-to-child heritability is evidenced on CHD of HR=2.949 (95% CI: 1.458–5.965, $p=2.14\times 10^{-3}$). It is worth
137 mentioning that a proportion of females born with CHD factually are not recommended for pregnancy. Among
138 148 expectant mothers with CHD from the ZEBRA maternity cohort, 14 (9.5%) neonates were also diagnosed
139 with CHD, whereas the incidence among the whole cohort was low at 1.2%. This finding is consistent with
140 previous studies²⁶, and the high conditional incidence rate suggests a potential genetic basis for CHD
141 development, although current medical research has yet to fully confirmed the aetiology. As a result, we firmly
142 endorse the implementation of genetic screening among eligible pregnant women with prenatally diagnosed
143 CHD in order to ascertain whether genetic anomalies underlie the congenital defects, since the probability of
144 developing CHD on offspring might be substantially amplified if there are genetic effects, warranting
145 preconception proactive attention in advance.

146 Maternal vascular diseases exert a greater impact than cardiac defects on the risk of APOs (HR=3.746, 95%
147 CI: 3.445–4.073, $p=3.71\times 10^{-33}$), but maternal blood disorders seem not to significantly affect neonates,
148 excluding anaemia which increases the risk of haematological conditions (HR=1.170, 95% CI: 1.072–1.277,
149 $p=5.01\times 10^{-4}$). Besides neonatal cardiovascular diseases, maternal vascular and blood anomalies also pose risks
150 on other birth defects (Fig. 3). While the majority of neonates (88.5%, 5,842 out of 6,604) did not exhibit
151 congenital cardiovascular anomalies at birth, a significant proportion of premature and low birth weight infants
152 could potentially encounter growth and developmental impairments, culminating in high medical burden and
153 unforeseeable compromise in long-term quality of life²⁷⁻²⁹. This reiterates the significance of extending
154 specialised medical vigilance towards pregnant individuals afflicted by cardiovascular conditions.

155 Maternal CVDs also modify the risks of APOs from environmental exposures. Pregnant participants with
156 CVDs manifest a 5.2% (4.9–5.5%, $p=1.85\times 10^{-39}$) higher risk regarding the PM_{2.5}-APOs association compared
157 to the control group without cardiovascular symptoms, and 2.2% higher (2.1–2.4%, $p=5.20\times 10^{-29}$) on O₃-APOs
158 risk association. The effect modification of maternal CVDs on greenness-APOs risk association is more
159 pronounced, thereby attenuating the significance of protective effects in certain subtypes of APOs. Fig. 4
160 provides a comprehensive overview of the intricate effect modifications associated with distinct subtypes of
161 maternal cardiovascular diseases. It demonstrates that maternal CVDs during pregnancy accentuate the
162 susceptibility of predisposed individuals to the adverse impacts of air pollution. For reproductive-age women
163 with primary cardiovascular symptoms, proactive strategies such as minimising exposure to ambient air
164 pollution or actively engaging with green spaces in the preconception period (i.e. at least 6–12 months prior to
165 pregnancy) are advised to attenuate the occurrence of perinatal anomalies.

166

167 **Mediation of maternal CVDs between environmental exposure and APOs**

168 From the perspective of a directed acyclic graph, maternal CVDs serve as mediators between environmental
169 exposure and APOs (as illustrated in [Extended Data Table 4](#)). Environmental exposures not only pose direct
170 impacts on APOs, but can also exert indirect effects through maternal CVDs. The indirect effects from PM_{2.5},
171 O₃, and greenness exposure on APOs account for 44.4%, 34.4%, and 23.1% of total effects, respectively,
172 underscoring the significance of mediation by maternal pregnancy-induced CVDs, particularly concerning the
173 risk associations with PM_{2.5} exposure. Environmental exposures still predominantly exert direct effects on APOs,
174 affirming the rationale for prior studies directly investigating the risk associations between environmental
175 exposures and APOs. Indirect effects are most pronounced in PTB and LBW, especially in the impacts on PTB
176 from PM_{2.5} exposure where the indirect effect constitutes over half proportion (50.9%). However, the impact of
177 greenness exposure on APOs is minimally mediated by maternal pregnancy-induced CVDs, particularly in the
178 context of LBW (12.1%) and neonatal CVDs (13.2%). These findings reveal a plausible causality pathway in
179 reproductive epidemiology, where environmental exposure influences maternal cardiovascular function,
180 subsequently impacting pregnancy outcomes. This also suggests the feasibility of reducing the risk of APOs by
181 controlling maternal cardiovascular disorders.

182 183 **Ensemble-learning-based risk prediction and self-directed assessment**

184 In virtue of the 60 characteristics collected on the ZEBRA maternity cohort participants together with the 40
185 environmental exposure tracking metrics, we developed an ensemble deep learning-based prediction model
186 ([Extended Data Fig. 4](#)) to forecast the incidence risk of APOs. Since early-stage risk prediction is a continuing
187 pursuit in clinical predictive models (i.e. predicting the occurrence of APOs during the third trimester lacks
188 practical value in timely clinical alert), predictive factors were selected upon the availability before the end of

189 the second trimester. The diagnosis of maternal primary and pregnancy-induced cardiovascular symptoms
190 significantly enhances the risk predictive performances of APOs. Without the cardiovascular diagnoses,
191 sensitivity (i.e. true positive predictions out of all cases) of APOs is limited to 59.7% (false positive rate,
192 FPR=0.8%, area under the ROC curve, AUC=0.899). However, cardiovascular diagnoses reinforce the
193 sensitivity of APOs to 80.8% (FPR=0.4%, AUC=0.956). For the collection of obstetric complications (i.e. PTB
194 and term LBW) and neonatal cardiorespiratory disorders (i.e. CHD, congenital cardiovascular defects, and
195 neonatal asphyxia), the ensemble learning prediction framework achieved accuracies as high as 97.9%
196 (sensitivity=82.8%, FPR=0.1%) and 98.5% (sensitivity=90.6%, FPR<0.1%), respectively. This exemplifies the
197 potential practical value of AI-assisted algorithmic risk prediction models for clinical application in obstetrics
198 and gynaecology, while also reaffirming the paramountcy of monitoring for maternal cardiovascular symptoms
199 during pregnancy.

200 Nevertheless, the confidentiality of maternal medical records and the restricted interpretability inherent in
201 machine learning algorithms preclude the feasibility of deploying ensemble-learning-based risk prediction
202 frameworks for public utilisation. To cope with this challenge, we developed a self-directed risk assessment
203 questionnaire (Supplementary Table S1) aimed at rapidly estimating APO risk scores by assigning weights to
204 factors easily collectable via self-reporting (Supplementary Table S2), followed by linear summation. Overall,
205 environmental exposures account for approximately 40% of the scoring weights, socioeconomic factors account
206 for 10%, while maternal medical and surgical history account for 50%, among which cardiovascular diseases
207 constitute 40%, comparable to environmental factors (Extended Data Figs. 5–6). While the predictive accuracy
208 of the self-directed form (88.2% for obstetric anomalies, 91.9% for neonatal cardiovascular defects) cannot fully
209 emulate the ensemble-learning-based models, its notable high sensitivities (75.7% for obstetric anomalies, 85.2%

210 for neonatal cardiovascular defects) still retain practical significance, inspiring future in-depth exploration and
211 optimisation.

212

213 **DISCUSSION**

214 To the best of our knowledge, this is the first Chinese prospective maternity cohort study comprehensively
215 exploring the risk patterns of maternal and neonatal cardiovascular abnormalities during the perinatal period,
216 providing robust primary epidemiological evidence specifically for the East Asian population. Our current study
217 has four major merits. Firstly, we conducted comprehensive analyses on the risk patterns of maternal and
218 neonatal cardiovascular abnormalities, including various subcategories such as cardiac abnormalities, vascular
219 diseases, and haematological disorders, expanding beyond the conventional focuses merely on stillbirth, PTB,
220 and term LBW^{30,31}. Secondly, we assessed the health hazards of multiple environmental factors individually as
221 well as investigated the inter-factor synergistic and antagonistic effects, as literature rarely explores the
222 interactions among risk factors. Thirdly, we quantified the mediating effect of maternal pregnancy-induced
223 cardiovascular complications on the risk association between environmental exposures and APOs, emphasising
224 the importance of protecting the pregnant female as a vulnerable population. Finally, we propose the feasibility
225 of using machine learning frameworks for early-stage APO prediction, based on which the designed self-
226 assessment form demonstrates a methodological innovation.

227 The APO risk self-assessment form we developed according to epidemiological evidence pioneers a new
228 method for practical public health research, that of translating sophisticated machine-learning-based algorithms
229 into interpretable linear algebra for intuitive weighting of risk factors and convenient risk prediction. We
230 encourage more maternity cohort studies to optimise and upgrade our risk prediction models by reporting
231 potential cross-region heterogeneity, so as to strengthen the population generalisability. When filling the self-

232 assessment form, the environmental exposure levels unavailable for self-report can easily be obtained by cloud
233 matching the residential location and conception dates of the pregnant women with the spatiotemporal resolved
234 environmental records. Therefore, to enable large-scale clinical implementation of APO risk prediction models,
235 the establishment of real-time environmental tracking databases encompassing multiple environmental factors
236 is indispensable.

237 With an increasing body of epidemiological research (including our current study) revealing the beneficial
238 impacts of greenness exposure on population health, we suggest policy recommendations for nature-based
239 interventions consider: i) enhancing residential greenery, such as the implementation of green roofs and the
240 creation of green streetscapes, particularly in areas heavily affected by air pollution; ii) enhancing the quality
241 of existing parks and gardens, which could involve improving the greenness accessibility, increasing the density
242 of vegetation, and providing outdoor fitness facilities; iii) advising pregnant women with cardiovascular
243 diseases to engage in outdoor nature-based activities as part of their treatment plans; iv) expanding the range of
244 community-led activities within green spaces, which could include initiatives like urban farms and retreat
245 centres; and v) promoting awareness and encouraging participation in wilderness programs, ecotherapy, and the
246 practice of forest bathing³².

247 The self-assessment form designed from the ZEBRA prospective maternity cohort will be continuously
248 calibrated and updated as more participants are recruited and more risk factors are collected, which will be
249 shared on MedRxiv preprint platforms. The intended value of the risk assessment form is enabling the pregnant
250 women to understand her pregnancy risks based on various exposures. Entrusting the complete authority of risk
251 assessment to authoritative institutions may give rise to corruption issues, like healthcare institutions
252 categorising more low-risk pregnant women as high-risk individuals for financial benefits. Empowering
253 pregnant women to be the sword-holders of their own health can result in pre-emptive action which then

254 effectively prevents such unnecessary medical overtreatment.

255 Our study has several limitations that cannot be addressed at this current stage. Firstly, ZEBRA has not yet
256 collected any genetic information from the cohort participants, which theoretically is the gold standard for
257 distinguishing endogenous differences among populations. Given that genetic markers can often substantially
258 improve the predictive accuracy of disease occurrence³³, the ZEBRA team plans to collect genetic material of
259 the maternity cohort in the near future. Secondly, we have not measured any internal exposure of organic
260 pollutants such as pesticide residuals, phthalates, and endocrine-disrupting chemicals. However, we have
261 initiated pilot studies to determine the exposure doses of pregnant women and fetuses^{34,35}, with the aim of
262 expanding to the entire cohort once the technology matures. Last but not the least, our findings may not be
263 representative of the entire Chinese population due to sparse participant enrolment in Northwest China.
264 Therefore, readers should use our results cautiously, particularly when generalising to broader populations, and
265 hence meta-analyses and heterogeneity tests are strongly recommended when more relevant studies come out.
266 It is optimistically hoped that multi-centre collaborative communities will be established soon to enhance the
267 representativeness and generalisability of epidemiological findings.

268

269 **METHODS**

270 **Cohort participant recruitment**

271 From 1 January 2017 to 31 December 2022, ZEBRA maternity cohort enrolled a total of 131,155 parturient women. In
272 addition to the 6,237 participants from the pilot study conducted during 2013–2016³⁶, the full cohort now includes a total
273 of 137,392 participants. Although the host institution (Women’s Hospital, Zhejiang University School of Medicine) of
274 ZEBRA is based in Zhejiang Province, it serves as one of the top three flagship maternity and child hospitals in China,
275 attracting expectant mothers from across the nation. Particularly, when pregnant women encounter complex medical
276 conditions during pregnancy and their local healthcare facilities are unable to provide definitive care, they are transferred
277 to Zhejiang, thus joining the ZEBRA maternity cohort. Among the current enrolled cohort participants, 122,440 (89.1%)

278 were from Zhejiang, while 14,952 (10.9%) were from other provinces, making the ZEBRA maternity cohort a nationwide
279 cohort in terms of its scale.

280 After excluding participants with missing records for more than 20% of the variables (N=16,091), those with
281 gestational ages less than 20 or greater than 44 weeks (N=39), and pregnant women aged under 18 or over 45 years (N=172),
282 our analyses cover 121,090 pregnant women in total (Fig. 1). There were 124,025 live births, with 116,610 singletons,
283 3,619 pairs of twins, and 59 sets of triplets, after excluding 802 stillbirths.

284 ZEBRA maternity cohort collects comprehensive sociodemographic and behavioural features from pregnant women
285 residing in Zhejiang Province, China throughout the study period. Trained obstetric nurses conduct questionnaire-based
286 face-to-face interviews to record residential address, household registration, ethnicity, education attainment level, smoking
287 habits (active and second-hand), alcohol history, age at current delivery, gravidity, parity, and date of the last menstrual
288 period. Physical examinations are performed to measure the height and weight of the women before conception and at the
289 end of the second trimester.

290 Gestational ages are primarily determined through ultrasound examinations conducted in the first or second trimester.
291 In cases where ultrasound records are unavailable, the date of the last menstrual period (LMP) is used as an alternative
292 method to estimate the gestational age. In this study, ultrasound examination was used for 98.4% (N=119,153) of the
293 pregnant women to determine gestational age, while the remaining 1.6% (N=1,937) relied on the date of the last menstrual
294 period.

295 To ensure data accuracy, the medical history of the pregnant women is cross-referenced with the hospital's medical
296 records using unique medical IDs. Face-to-face consultations with obstetricians are conducted to verify the referenced
297 information. The cohort profile provides a comprehensive overview of the collected information and the processes
298 involved²⁴. Stringent quality control measures are in place, including regular training for healthcare and medical personnel,
299 standardised medical examinations, use of uniform-standard physiological and biochemical measurement equipment, and
300 double validation of questionnaire-based interviews.

301 **Outcome definition**

302 In our study, we identified primary (referring to the non-pregnancy-cause diseases throughout all ZEBRA-based studies
303 no matter congenital or acquired) and pregnancy-induced (a type of secondary) cardiovascular complications in all enrolled
304 cohort participants. Primary maternal CVDs include primary hypertension (ICD10: I10), pulmonary hypertension (I27),
305 congenital heart diseases (Q20–Q28), heart failure (I30–I45), and cardiac arrhythmias (I47–I49). Haematological disorders
306 (D50–D89) include anaemia (including nutritional anaemia, D50–D53, and haemolytic anaemia, D55–D59), and lymphatic

307 system abnormalities (D76). Pregnancy-induced CVDs include gestational hypertension (O13) and pre-eclampsia (O14).
308 The hierarchical classification and occurrence rates of the studied diseases can be found in [Extended Data Fig. 1](#).

309 Adverse pregnancy outcomes (APOs) include stillbirth, obstetric anomalies, neonatal cardiovascular diseases,
310 neonatal haematological diseases, and neonatal respiratory disorders. Obstetric anomalies include preterm birth (PTB, O60)
311 and term low birth weight (LBW, P07.1). PTB is defined as gestation less than 37 weeks, and very preterm birth (VPTB)
312 is defined as gestation less than 32 weeks. Term LBW is defined as birth weight less than 2500 grams on the term-birth
313 neonates, while very low birth weight (VLBW) refers to term-birth neonates with birth weight less than 1500 grams. When
314 evaluating risk associations related to obstetric complications, we only considered singleton births.

315 Neonatal cardiovascular complications mainly include cardiac disorders and pulmonary hypertension. Cardiac
316 abnormalities encompass congenital heart diseases and cardiac arrhythmias, while haematological abnormalities include
317 coagulation defects, purpura, and other haemorrhagic disorders (D65–D69). ZEBRA provides accurate diagnoses for
318 congenital heart diseases in newborns, with subcategories mainly including transposition of great vessels, ventricular septal
319 defects (Q21.0), atrial septal defects (Q21.1), tetralogy of Fallot (Q21.3), patent ductus arteriosus (Q25.0), coarctation of
320 the aorta (Q25.1–Q25.3), and pulmonary stenosis (Q25.5–Q25.6). In this study, we do not further classify maternal
321 congenital heart diseases into subcategories. Neonatal respiratory abnormalities include birth asphyxia (P21) and
322 respiratory distress (P22). The classification and incidence rates of the analysed diseases can be found in [Extended Data](#)
323 [Fig. 2](#).

324 It should be clarified that throughout our present analyses, APOs are defined as all the conditions mentioned above.
325 The cases of PTB and term LBW are counted per pregnancy (equivalent to the number of women with abnormal pregnancy
326 outcomes), as multiple births are often associated with medically induced premature delivery.

327 **Environmental exposure assessment**

328 In this study, we evaluated maternal exposure to three major environmental factors: ambient PM_{2.5}, O₃, and green space.
329 We tracked the historical concentrations of ambient PM_{2.5} (in daily average) and O₃ (in daily maximum 8-hour moving
330 average) during the period from 2013 to 2022 using the TAP (Tracking Air Pollution) database^{37,38}. The TAP database
331 integrates various data sources, including *in situ* observations, numerical simulations from chemical transport models,
332 satellite-based remote sensing measurements, and land cover information. These data are combined using an ensemble
333 machine learning framework to provide highly accurate daily concentration estimates. The spatial resolution of PM_{2.5} data
334 is 1×1 km², while the resolution of O₃ data is 10×10 km². Green spaces are assessed using the Enhanced Vegetation Index
335 (EVI) provided by MODIS Vegetation Index Products³⁹. The spatial resolution of EVI data is 0.5×0.5 km², and

336 measurements are available every 16 days. Maternal exposures were averaged over an 18-month period, from one year
337 before pregnancy until the end of the second trimester (the 6th month since conception).

338 We also tracked individual-level exposure to three additional ambient air pollutants, NO₂, SO₂, and CO, as covariates
339 for confounding adjustment. These exposure measurements were obtained from other well-established datasets⁴⁰⁻⁴². Due
340 to the limitations of historical data availability, the tracking of these air pollutants is only available until December 2020.
341 To overcome this limitation, we extrapolated the pollution concentrations from the corresponding day in 2020 to serve as
342 proxies for the years 2021–2022. Considering the temporal extrapolation and the absence of significant risk associations,
343 the analysis results for these three air pollutants are not presented as the main findings. Furthermore, we accounted for the
344 influence of temperature exposure by controlling for the impact of heat index, which is a humidity-calibrated temperature
345 measure that better captures the perceived sensation of temperature exposure^{43,44}. We quantified two heat exposure metrics:
346 daily mean heat index and daily maximum heat index. Daily ambient air temperature data with a resolution of 0.1°×0.1°
347 and humidity data with a resolution of 0.25°×0.25° were retrieved from ECMWF Reanalysis v5 (ERA5) products⁴⁵.

348 **Statistical analyses**

349 **Risk association quantification.** Extended Cox proportional hazard regression models with time-varying variables were
350 applied to investigate i) the hazard ratios (HR) and 95% confidence intervals (CI) of environmental exposures on maternal
351 pregnancy-induced cardiovascular complications; ii) the HR of environmental exposures on APO and subtypes; and iii)
352 the HR of maternal cardiovascular complications (both primary and pregnancy-induced) on APO and subtypes. Assuming
353 consistent HRs over time, the temporal granularity was delineated on a weekly basis, spanning from the preconception
354 year to the occurrence of the specified health outcomes rather than the conventional gestational age.

355 The crude regression models solely examine the direct association between the studied risk factor, denoted as X_0 , and
356 the occurrence of the target health outcomes (Equation 1). The fully adjusted regression model considers other risk factors
357 and potential confounders (such as other environmental exposure factors, socioeconomic characteristics, medical history,
358 and the sex of neonates, denoted as X_i) (Equation 2). Additionally, interaction terms between the studied pairs of factors
359 are included in the extended model to assess the effect modification (EM) between risk factors (Equation 3).

$$h(t) = h_0(t)e^{\beta_0 X_0} \quad \text{(Equation 1)}$$

$$h(t) = h_0(t)e^{\beta_0 X_0 + \sum \beta_i X_i} \quad \text{(Equation 2)}$$

$$h(t) = h_0(t)e^{\beta_0 X_0 + \sum \beta_i X_i + \sum_{i \neq j} \gamma_{ij} X_i X_j} \quad \text{(Equation 3)}$$

360 The coefficients β_i obtained from the Cox regression models accompanied with standard errors (SE_i) are transformed
361 to HR following Equation 4, where Δx represents a unit incremental exposure in the risk factor. For binary variables, Δx is

362 defined as 1, and transformed HRs indicate the increased risk comparing with and without the corresponding risk factor.
363 For environmental exposures, Δx is defined as 10- $\mu\text{g}/\text{m}^3$, 10-ppb, and 0.1-EVI increments of $\text{PM}_{2.5}$, O_3 , and greenness,
364 respectively. The transformed HR signifies the increased risk associated with each 10 $\mu\text{g}/\text{m}^3$ increment in $\text{PM}_{2.5}$ exposure;
365 the interpretations for O_3 and green space follow the similar way. EM is calculated by Equation 5, which can be interpreted
366 as the alteration of one exposure factor on the association between another environmental exposure and risk of the studied
367 adverse health outcome, under the specified increment of environmental exposure.

$$HR = e^{\beta\Delta x} \quad (\text{Equation 4})$$

$$EM = e^{\gamma_{ij}\Delta x_i\Delta x_j} \quad (\text{Equation 5})$$

368 **Direct and indirect effect evaluation.** Given that maternal CVDs are positioned along the causal pathway between
369 environmental exposure and APOs, maternal CVDs are considered mediators from a biostatistical perspective. Multivariate
370 Cox regression models were applied to examine the risk associations quantified in log-transformed HR between i)
371 environmental exposure and APOs having adjusted for all the other covariates except the maternal CVDs, marked as β_1 ,
372 and ii) environmental exposure and APOs having adjusted for all studied covariates including the maternal CVDs, marked
373 as β_2 . β_2 represents the direct effects on APOs from environmental exposure, while $\beta_1 - \beta_2$ indicates the indirect effects.
374 Analysed maternal CVDs as mediators include pregnancy-induced hypertension and preeclampsia, and APOs include PTB,
375 LBW, and neonatal CVDs.

376 **Exposure-response trend determination.** We employed restricted cubic spline regression with no more than 4 degrees of
377 freedom to investigate the risk association trends of maternal $\text{PM}_{2.5}$, O_3 , and greenness exposures on pregnancy-induced
378 cardiovascular diseases (including hypertension), obstetric adverse pregnancy outcomes (including preterm birth and term
379 low birth weight), and neonatal cardiovascular diseases. All covariates considered in the multivariate models are retained.
380 The risk thresholds for the three environmental factors are defined as the 5th percentile of exposure levels among all cohort
381 participants⁴⁶.

382 **Inter-group comparison of risk associations.** We performed Cox regressions on participants with specific cardiovascular
383 symptoms (including subcategories) and those without specific cardiovascular symptoms to obtain grouped HRs with 95%
384 CIs. Subsequently, we conducted Levene's tests on the paired log-transformed HRs (i.e. $\beta_i \pm SE_i$) for each group to
385 determine the homogeneity of variances between SE_1 and SE_2 , in order to decide whether to use a homoscedastic or
386 heteroscedastic *t*-test. Statistical significance (*p*-value) was defined as $\alpha < 0.05$ for two-sided tests; the *t*-test *p*-values
387 (p_{hom} or p_{het}) are calculated following Equations 6–11, where n_1 and n_2 represent the sample sizes of the two
388 groups; *df* stands for degrees of freedom; s_1 and s_2 for the standard deviations of either group; s_p for the pooled standard
389 deviation; t_{hom} and t_{het} for homoscedastic and heteroscedastic *t*-statistics, respectively; Γ for gamma distribution; p_i for

390 p -values of homoscedastic and heteroscedastic t -tests, respectively.

$$df_{hom} = n_1 + n_2 - 2 \quad (\text{Equation 6})$$

$$df_{het} = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}} \quad (\text{Equation 7})$$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{df_{hom}} \quad (\text{Equation 8})$$

$$t_{hom} = \frac{\beta_1 - \beta_2}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}} \quad (\text{Equation 9})$$

$$t_{het} = \frac{\beta_1 - \beta_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (\text{Equation 10})$$

$$p_i = \int_{|t_i|}^{\infty} \frac{\Gamma\left(\frac{df_i + 1}{2}\right)}{\sqrt{df_i \pi} \Gamma\left(\frac{df_i}{2}\right)} \left(1 + \frac{x^2}{df_i}\right)^{-\frac{df_i + 1}{2}} dx \quad (\text{Equation 11})$$

391 All statistical analyses in this study were performed in Stata 17. The prediction model was performed via the “*scikit-*
392 *learn*” (version 1.2.0) and “*xgboost*” package (version 1.6.2) in Python 3.8.0. Computations were supported by JASMIN
393 supercomputer.

394 **Ensemble machine learning algorithm for risk prediction**

395 We integrated 10 socioeconomic features, 30 environmental exposure indicators, 10 items for pregnancy-relevant maternal
396 cardiovascular diagnoses, 21 items for obstetric disease diagnoses, and 29 items of other aspects of medical history to
397 predict the risks of obstetric APOs and neonatal cardiovascular diseases by virtue of ensemble learning. Initially, we
398 employed 6 classical base learners, as i) linear logistic regression classifier, ii) decision tree classifier, iii) random forest
399 classifier, iv) extra-trees classifier, v) bootstrap aggregating (bagging) classifier, and vi) gradient boosting classifier, to
400 construct the supervised training models and generate predicted probability scores of the studied adverse health outcomes.
401 Subsequently, we used a fully connected multi-layer perceptron classifier to integrate the original predictive factors and
402 the risk probability scores obtained from the base learners to build a final predictive model for the occurrence of the two
403 categories of APOs. The multi-layer perceptron classifier consists of a fully connected artificial neural network with 5
404 hidden layers, each composed of 256, 256, 128, 64, and 32 nodes, respectively (refer to Extended Data Fig. 4 for the
405 algorithm structure). The hyperparameters (number of hidden layers and nodes) of the multi-layer perceptron are
406 determined by learning curves.

407 We evaluated the predictive performance of the final multi-layer perceptron classifier by randomly selecting 80% of

408 the samples for model training and performed 10-fold cross-validation tests; the remaining 20% of the samples (N=24,218)
409 were used for external validation of the model. Evaluation also included assessing the sensitivity, false positive rate, and
410 the overall power of the prediction model (area under the ROC curve, AUC).

411 **Design of self-directed APO risk assessment form**

412 We utilised ensemble learning algorithms to forecast the probability of risk events transpiring, employing the prediction
413 score as the dependent variable. We consider all discretised risk factors as independent variables and ascertain the weighting
414 coefficients through linear regression. Continuous variables were discretised using quartiles (e.g. environmental exposure
415 levels) or predetermined segmenting criteria (e.g. BMI). The procedure for determining the weights of the risk factors can
416 be outlined as follows:

- 417 1) Perform a linear regression with intercept β on the target score P against the discrete risk factors \mathbf{X}_D , generating
418 an initial weight matrix \mathbf{A}_0 , in terms of $P = \mathbf{A}_0\mathbf{X}_D + \beta$.
- 419 2) For risk factors with negative weights, reverse the sequence of the corresponding discrete labels $\tilde{\mathbf{X}}_D$, and
420 transform the weight matrix \mathbf{A}_0 into a non-negative matrix, $\mathbf{A}_0^+ = |\mathbf{A}_0|$.
- 421 3) Define $\mathbf{A}_1 = \mathbf{A}_0^+$, where subscript indicates the round of iteration.

422 **Iteration start:**

- 423 - Compute the weighted score excluding the intercept: $\hat{P} = \mathbf{A}_i\tilde{\mathbf{X}}_D$.
- 424 - Select the case-control overlapping samples, based on 95% interval (2.5–97.5th percentile) of the score
425 distribution for cases and controls.
- 426 - Perform a linear regression with intercept on sample set, generate a new weight matrix, \mathbf{A}' , for the current
427 iteration.
- 428 - Adjust the weight matrix: $\mathbf{A}_{i+1} = \frac{\mathbf{A}_i + \mathbf{A}'}{i+1}$.
- 429 - Check if either i) $|\mathbf{A}_{i+1} - \mathbf{A}_i| < 0.001$, or ii) the number of overlap samples is less than 5% of the total
430 number of cases.

431 **Iteration ends.**

- 432 4) Calculate the sum of the maximum weights for each factor under the adjusted weight matrix after iteration, \mathbf{A}_{i+1} ,
433 and rescale it to 100, yielding the final weights, \mathbf{A}_s .
- 434 5) Compute the fully optimised risk scores for all pregnant women in the cohort with rescaled weights: $\hat{P} = \mathbf{A}_s\tilde{\mathbf{X}}_D$.
435 Define the median score of the overlap interval for cases and controls as the risk forecast threshold.
- 436 6) Complete the tabulation (see [Supplementary Table S1](#) for a sample table).

437 We conducted external validation for the self-directed APO risk prediction form on 5,217 pregnant women recruited
438 between 1 January and 31 May 2023. The de-identified questionnaires of two pregnant women with the highest predicted
439 risk scores of obstetric APOs and neonatal cardiovascular defects are enclosed in [Supplementary Tables S3–S4](#) as samples.

440 **Sensitive analysis**

441 We conducted multiple sensitivity analyses to ensure the robustness of our results. These analyses include: 1) randomly
442 selecting 50% of the original samples for five times of statistical analysis ([Supplementary Tables S5–S6](#)); 2) using different
443 combinations of environmental exposures to adjust for potential confounders ([Supplementary Table S7](#)); and 3) dividing
444 the whole studied population into two subgroups, within Zhejiang and outside Zhejiang Province, so as to test the inter-
445 group heterogeneity of the estimated risk associations ([Supplementary Table S8](#)).

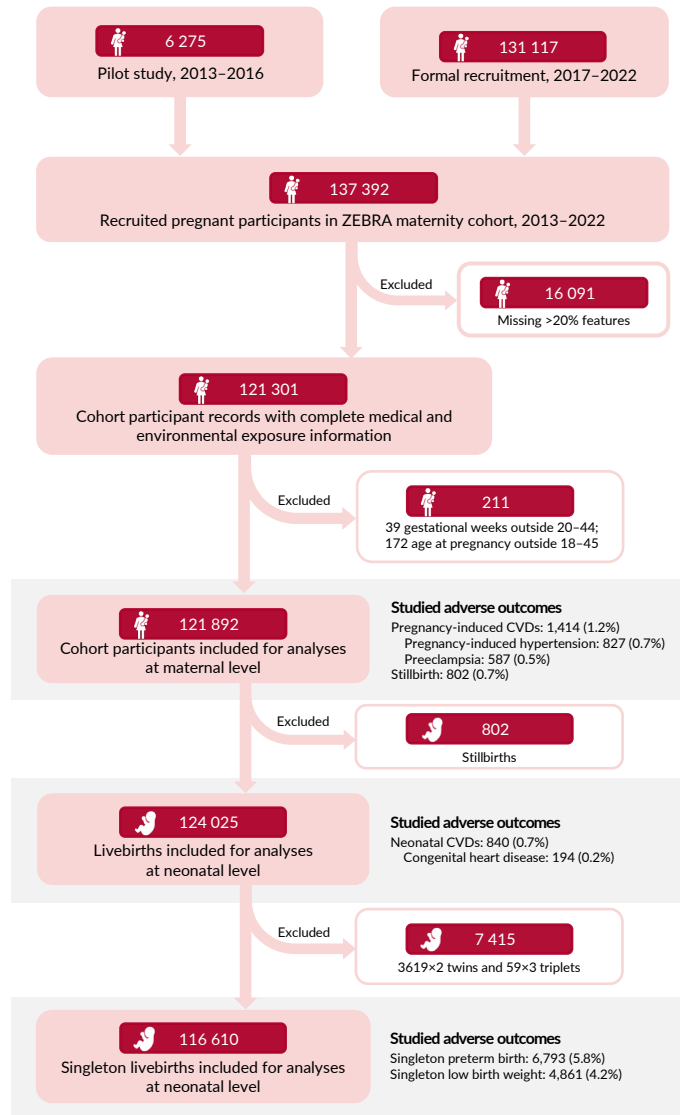
446 **Ethics Approval**

447 The current cohort-based study was approved by the ethics committee of the Women’s Hospital, Zhejiang University
448 School of Medicine (IRB-20220189-R) and we obtained written informed consent from all cohort participants upon
449 enrolment. As an original investigation providing first-hand epidemiological evidence, Strengthening the Reporting of
450 Observational Studies in Epidemiology (STROBE) guidelines for result reporting (see in [Supplementary Table S9](#)) was
451 strictly followed.

452

453 **TABLES AND FIGURES**

454



455

456

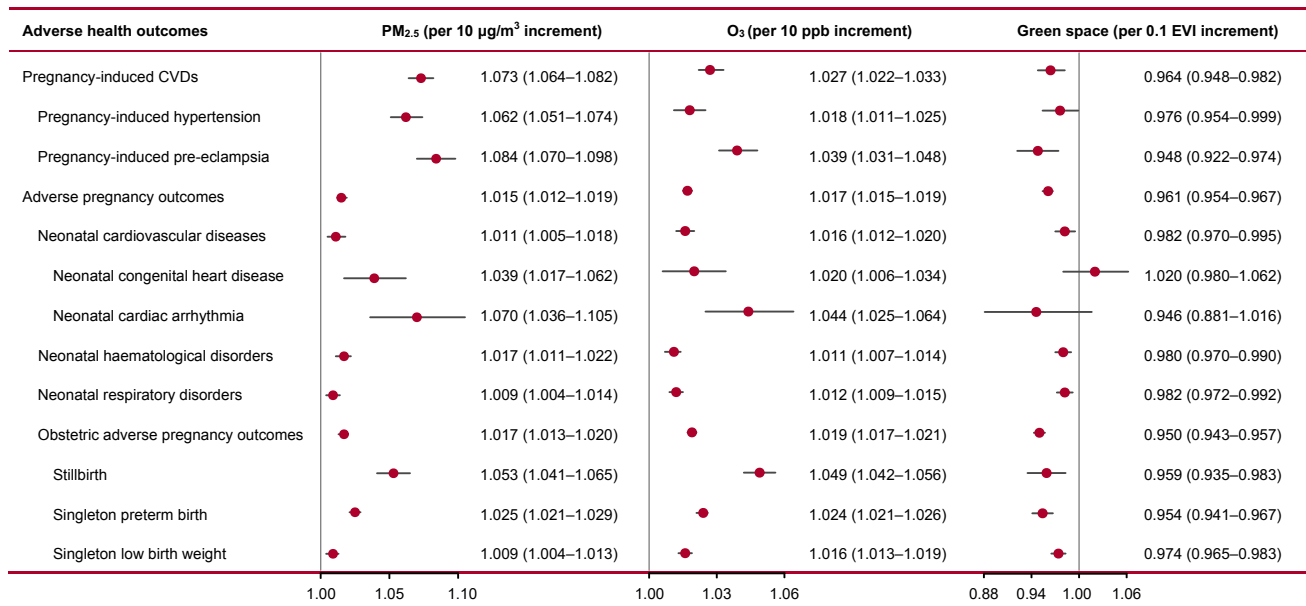
Fig. 1 | Recruitment flowchart of ZEBRA maternity cohort participants, 2013–2022.

457

ZEBRA maternity cohort comprises a comprehensive dataset, including sociodemographic characteristics, medical history prior to conception, diagnoses during pregnancy, obstetric-related diagnoses during the perinatal period, physiological and biochemical parameters obtained through routine blood and urine tests, and a retrospective record of environmental exposure tracing. Participants who changed their place of residence during the 18-month period for analysis (from one year before conception to the end of the second trimester) or had records missing >20% of studied variables were excluded from the analysis. Additionally, participants with gestational durations <20 or >44 weeks and aged <18 or >45 years were censored to eliminate potential effects on pregnancy outcomes. Shaded participants represent the cohort members included in designed analyses investigating three epidemiological aspects: i) the influence of risk factors on the occurrence of maternal pregnancy-induced cardiovascular diseases and stillbirths; ii) the relationship between risk factors and neonatal cardiovascular disorders; and iii) the association between risk factors and obstetric anomalies in singleton neonates.

467

468



469

470

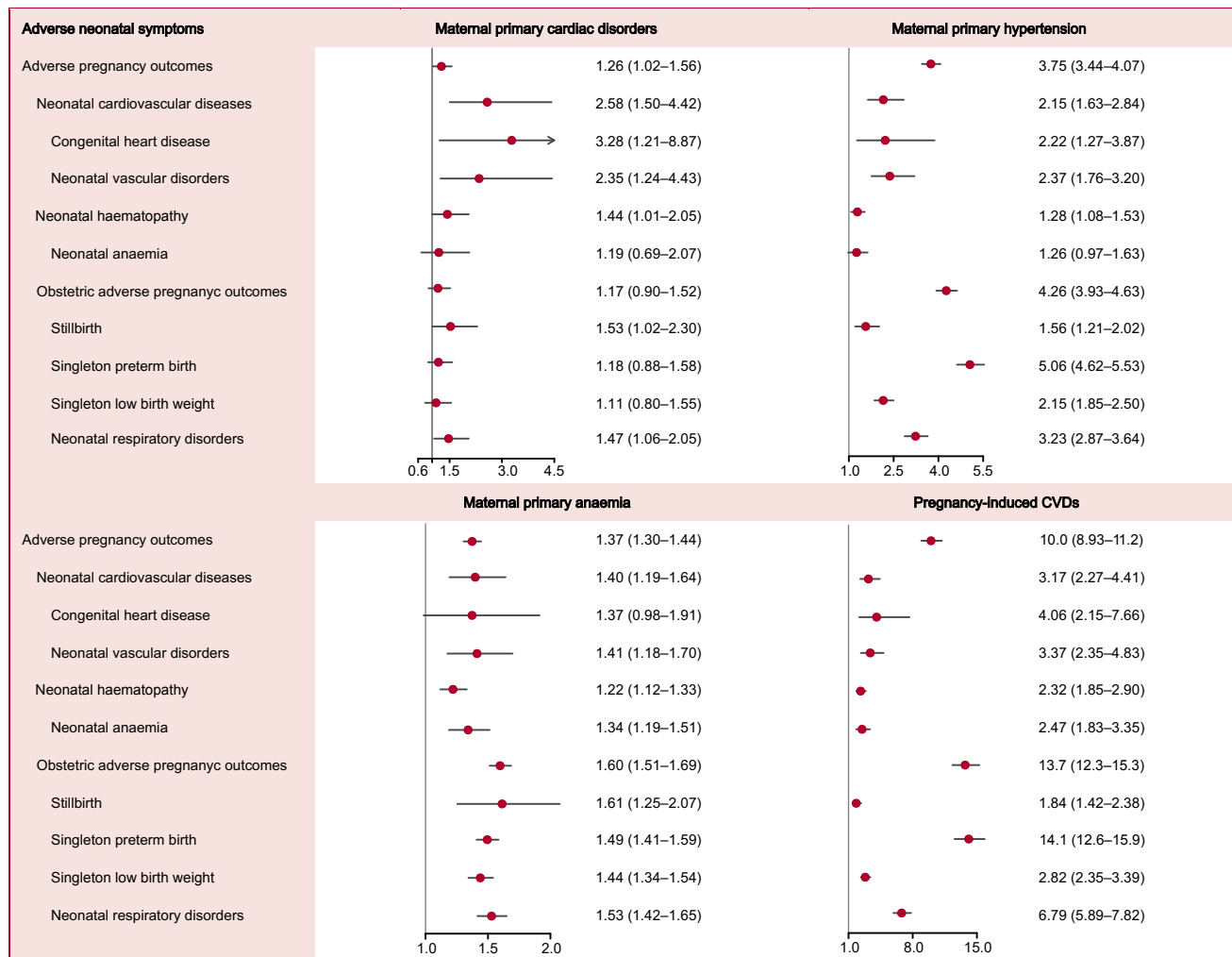
Fig. 2 | Risk association between environmental exposure and perinatal abnormalities.

471

Epidemiological analyses encompass a variety of adverse health outcomes, including maternal pregnancy-induced cardiovascular diseases (CVDs) (excluding primary CVDs), neonatal cardiovascular and haematological disorders, as well as stillbirth, singleton preterm birth, and singleton low birth weight. Environmental exposure tracking covers PM_{2.5}, O₃, and greenness, to which hazard ratios (HR) with 95% confidence intervals (CI) were estimated sufficiently having adjusted for covariates for 10-µg/m³, 10-ppb, and 0.1-EVI incremental exposure, respectively. Subcategories of adverse health outcomes that did not exhibit significant risk associations with three environmental exposures were not included in the forest-plot. The indentation degree of listed subcategories of adverse health outcomes represents hierarchical tiers.

478

479



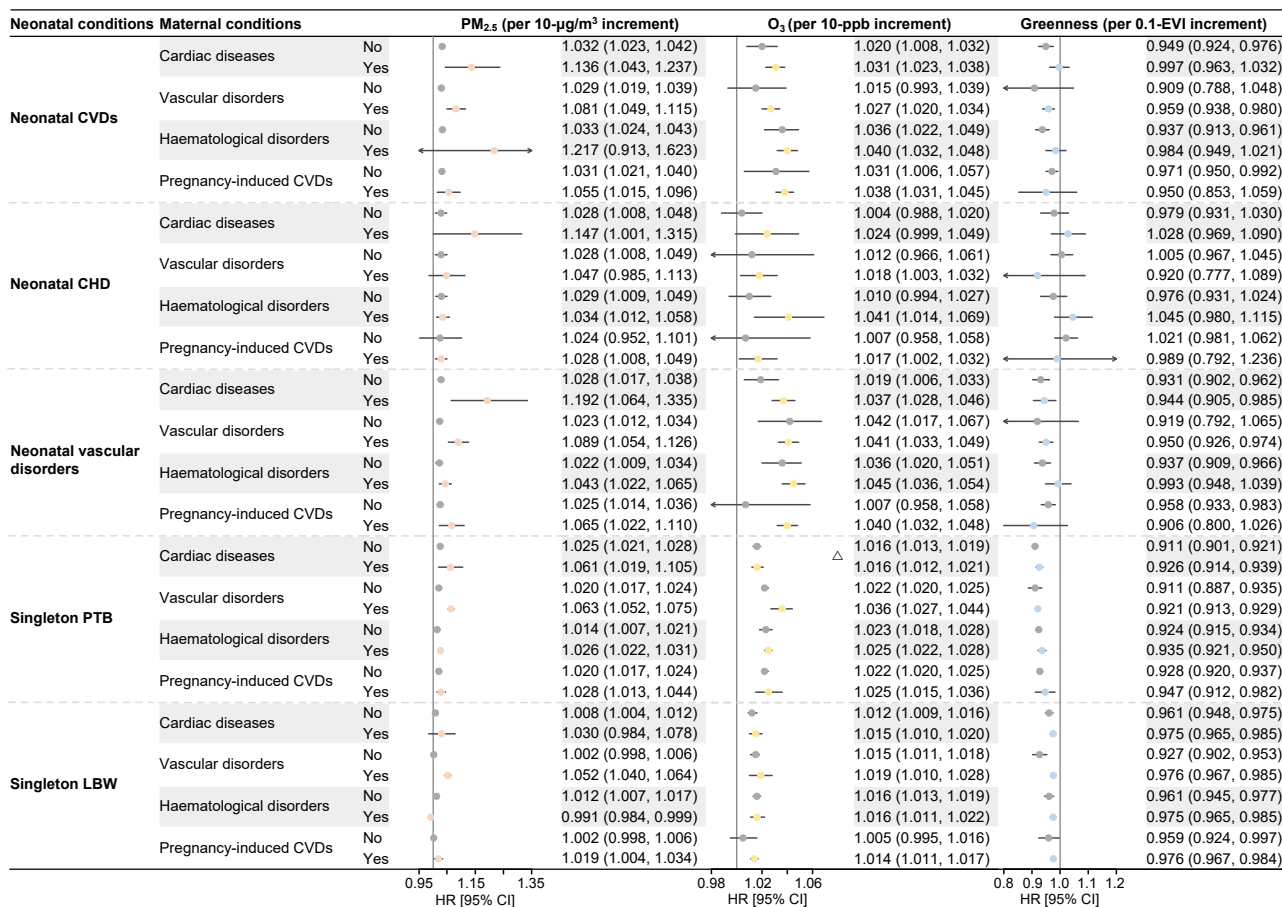
480

481 **Fig. 3 | Risk association between maternal cardiovascular and haematological complications and**
 482 **adverse pregnancy outcomes.**

483 Risk factors are defined as maternal CVDs, including primary cardiac disorders, primary hypertension, primary
 484 anaemia, and pregnancy-induced CVDs. Adverse health outcomes encompass obstetric anomalies (stillbirth,
 485 singleton PTB, singleton LBW, and neonatal respiratory disorders), as well as neonatal cardiovascular and
 486 haematological disorders. Multiple pregnancies are highly likely to be associated with medically indicated PTB and
 487 naturally occurring LBW, and thus are typically excluded from the analysis on investigating obstetric anomalies.
 488 Hazard ratios (HR) with 95% confidence intervals (CI) were estimated by Cox regression models having adjusted for
 489 covariates for 10- $\mu\text{g}/\text{m}^3$, 10-ppb, and 0.1-EVI incremental exposure, respectively.

490

491



492

493

494

495

496

497

498

499

500

501

502

Fig. 4 | Effect modification of maternal cardiovascular complications on the risk association between environmental exposure and adverse pregnancy outcomes.

Four effect modifiers, including maternal primary cardiac diseases, primary vascular disorders, primary haematological disorders, and pregnancy-induced CVDs, were examined. The presentations of estimated hazard ratios (HRs) are defined along three dimensions: neonatal adverse health condition (N), maternal primary and pregnancy-induced diseases (M), and environmental exposure (E), in which context HRs represent the strength of the risk association between E and N, with or without the presence of M. Within the M group, no significant difference at $\alpha=0.05$ between the two HR levels was indicated by Δ , while the unmarked groups signify that there is a significant difference in risks with or without M.

503 **REFERENCES**

504

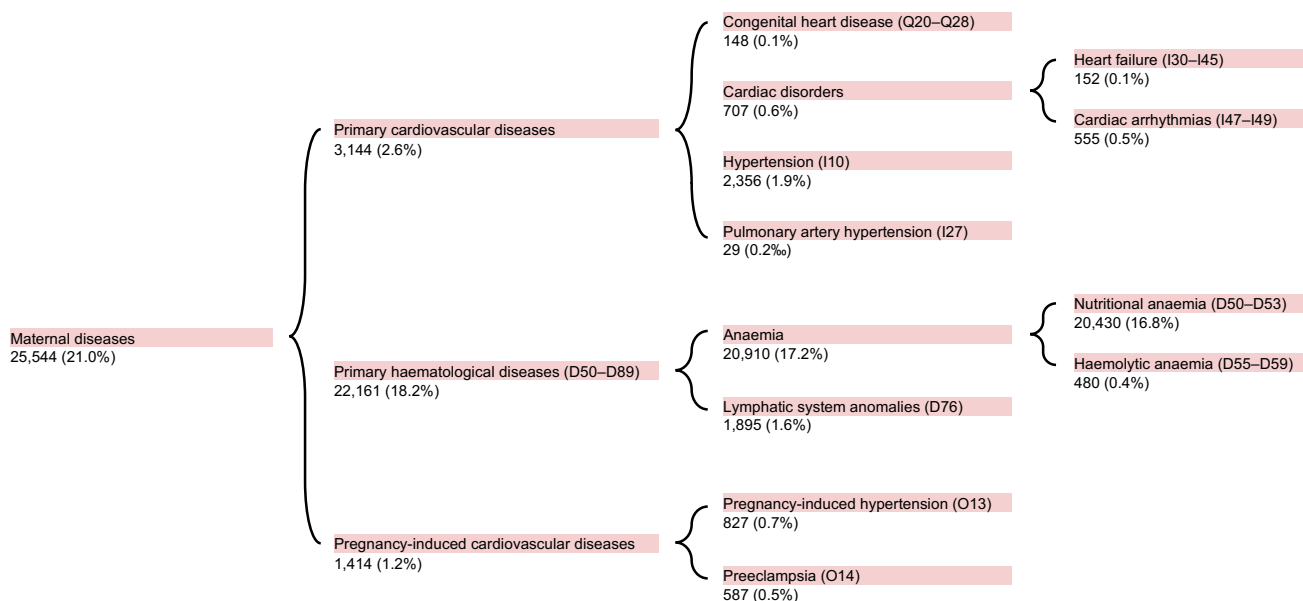
- 505 1. Tatum M. China's three-child policy. *Lancet* 2021; **397**(10291): 2238.
- 506 2. Sun HZ, Xiang Q, Xu S, et al. China's unwavering determination in protecting pregnancy and perinatal health. *The*
507 *Innovation* 2022; **3**(6): 100336.
- 508 3. World Health Organization. World health statistics 2023: monitoring health for the SDGs, sustainable development
509 goals: World Health Organization; 2023.
- 510 4. Kruk ME, Gage AD, Arsenault C, et al. High-quality health systems in the Sustainable Development Goals era: time
511 for a revolution. *Lancet Glob Health* 2018; **6**(11): e1196-e252.
- 512 5. Sun HZ, Yu P, Lan C, et al. Cohort-based long-term ozone exposure-associated mortality risks with adjusted
513 metrics: A systematic review and meta-analysis. *The Innovation* 2022; **3**(3): 100246.
- 514 6. Verhoeven JI, Allach Y, Vaartjes ICH, Klijn CJM, de Leeuw FE. Ambient air pollution and the risk of ischaemic and
515 haemorrhagic stroke. *Lancet Planet Health* 2021; **5**(8): e542-e52.
- 516 7. de Bont J, Jaganathan S, Dahlquist M, Persson A, Stafoggia M, Ljungman P. Ambient air pollution and
517 cardiovascular diseases: An umbrella review of systematic reviews and meta-analyses. *J Intern Med* 2022; **291**(6):
518 779-800.
- 519 8. Shah AS, Lee KK, McAllister DA, et al. Short term exposure to air pollution and stroke: systematic review and
520 meta-analysis. *BMJ* 2015; **350**: h1295.
- 521 9. Liu XX, Ma XL, Huang WZ, et al. Green space and cardiovascular disease: A systematic review with meta-analysis.
522 *Environ Pollut* 2022; **301**: 118990.
- 523 10. Sun HZ, Zhao J, Liu X, et al. Antagonism between ambient ozone increase and urbanization-oriented population
524 migration on Chinese cardiopulmonary mortality. *The Innovation* 2023; **4**(6): 100517.
- 525 11. Ren M, Wang Q, Zhao W, et al. Effects of extreme temperature on the risk of preterm birth in China: A population-
526 based multi-center cohort study. *Lancet Reg Health West Pac* 2022; **24**: 100496.
- 527 12. Zhang L, Shi S, Wu S, et al. Effects of greenness on preterm birth: A national longitudinal study of 3.7 million
528 singleton births. *The Innovation* 2022; **3**(3): 100241.
- 529 13. Chen J, Guo L, Liu H, et al. Modification effects of ambient temperature on associations of ambient ozone exposure
530 before and during pregnancy with adverse birth outcomes: A multicity study in China. *Environ Int* 2023; **172**:
531 107791.
- 532 14. Bekkar B, Pacheco S, Basu R, DeNicola N. Association of Air Pollution and Heat Exposure With Preterm Birth,
533 Low Birth Weight, and Stillbirth in the US: A Systematic Review. *JAMA Netw Open* 2020; **3**(6): e208243.
- 534 15. Kloog I. Air pollution, ambient temperature, green space and preterm birth. *Curr Opin Pediatr* 2019; **31**(2): 237-43.
- 535 16. Zhang LQ, Liu WW, Hou K, et al. Air pollution-induced missed abortion risk for pregnancies. *Nat Sustain* 2019;
536 **2**(11): 1011-7.
- 537 17. Veerbeek JH, Hermes W, Breimer AY, et al. Cardiovascular disease risk factors after early-onset preeclampsia, late-
538 onset preeclampsia, and pregnancy-induced hypertension. *Hypertension* 2015; **65**(3): 600-6.
- 539 18. Kurabayashi T, Mizunuma H, Kubota T, Kiyohara Y, Nagai K, Hayashi K. Pregnancy-induced hypertension is
540 associated with maternal history and a risk of cardiovascular disease in later life: Japanese cross-sectional study.
541 *Maturitas* 2013; **75**(3): 227-31.
- 542 19. Steegers EA, von Dadelszen P, Duvekot JJ, Pijnenborg R. Pre-eclampsia. *Lancet* 2010; **376**(9741): 631-44.
- 543 20. Udine ML, Evans F, Burns KM, Pearson GD, Kaltman JR. Geographical variation in infant mortality due to
544 congenital heart disease in the USA: a population-based cohort study. *Lancet Child Adolesc Health* 2021; **5**(7): 483-
545 90.
- 546 21. Gratz A, Hess J, Hager A. Self-estimated physical functioning poorly predicts actual exercise capacity in

- 547 adolescents and adults with congenital heart disease. *Eur Heart J* 2009; **30**(4): 497-504.
- 548 22. Ladouceur M, Benoit L, Radojevic J, et al. Pregnancy outcomes in patients with pulmonary arterial hypertension
549 associated with congenital heart disease. *Heart* 2017; **103**(4): 287-92.
- 550 23. Wang D, Zhang Y, Jiang Y, et al. Shanghai Preconception Cohort (SPCC) for the association of periconceptional
551 parental key nutritional factors with health outcomes of children with congenital heart disease: a cohort profile. *BMJ*
552 *open* 2019; **9**(11): e031076.
- 553 24. Sun HZ, Tang H, Xiang Q, et al. Cohort Profile: Zhejiang Environmental and Birth Health Research Alliance
554 (ZEBRA) Maternity Cohort. *medRxiv* 2023: 2023.02.21.23286173.
- 555 25. Sun HZ, Tang H, Fang J, et al. A Chinese longitudinal maternity cohort study (2013–2021) on intrahepatic
556 cholestasis phenotypes: Risk associations from environmental exposure to adverse pregnancy outcomes. *J Hazard*
557 *Mater* 2023; **463**: 132915.
- 558 26. van der Bom T, Zomer AC, Zwinderman AH, Meijboom FJ, Bouma BJ, Mulder BJ. The changing epidemiology of
559 congenital heart disease. *Nat Rev Cardiol* 2011; **8**(1): 50-60.
- 560 27. Linsell L, Malouf R, Morris J, Kurinczuk JJ, Marlow N. Prognostic Factors for Poor Cognitive Development in
561 Children Born Very Preterm or With Very Low Birth Weight: A Systematic Review. *JAMA Pediatr* 2015; **169**(12):
562 1162-72.
- 563 28. Luyckx VA, Bertram JF, Brenner BM, et al. Effect of fetal and child health on kidney development and long-term
564 risk of hypertension and kidney disease. *Lancet* 2013; **382**(9888): 273-83.
- 565 29. Blencowe H, Lee AC, Cousens S, et al. Preterm birth-associated neurodevelopmental impairment estimates at
566 regional and global levels for 2010. *Pediatr Res* 2013; **74**(Suppl 1): 17-34.
- 567 30. Xue T, Tong M, Li J, et al. Estimation of stillbirths attributable to ambient fine particles in 137 countries. *Nat*
568 *Commun* 2022; **13**(1): 6950.
- 569 31. Fleischer NL, Merialdi M, van Donkelaar A, et al. Outdoor air pollution, preterm birth, and low birth weight:
570 analysis of the world health organization global survey on maternal and perinatal health. *Environ Health Persp*
571 2014; **122**(4): 425-30.
- 572 32. Shanahan DF, Astell-Burt T, Barber EA, et al. Nature-Based Interventions for Improving Health and Wellbeing:
573 The Purpose, the People and the Outcomes. *Sports* 2019; **7**(6): 141.
- 574 33. Chatterjee N, Shi J, Garcia-Closas M. Developing and evaluating polygenic risk prediction models for stratified
575 disease prevention. *Nat Rev Genet* 2016; **17**(7): 392-406.
- 576 34. Chen F, Yin S, Kelly BC, Liu W. Chlorinated Polyfluoroalkyl Ether Sulfonic Acids in Matched Maternal, Cord, and
577 Placenta Samples: A Study of Transplacental Transfer. *Environ Sci Technol* 2017; **51**(11): 6387-94.
- 578 35. Liu Y, Li A, Buchanan S, Liu W. Exposure characteristics for congeners, isomers, and enantiomers of perfluoroalkyl
579 substances in mothers and infants. *Environ Int* 2020; **144**: 106012.
- 580 36. Sun Z, Yang L, Bai X, et al. Maternal ambient air pollution exposure with spatial-temporal variations and preterm
581 birth risk assessment during 2013-2017 in Zhejiang Province, China. *Environ Int* 2019; **133**(Pt B): 105242.
- 582 37. Geng G, Xiao Q, Liu S, et al. Tracking Air Pollution in China: Near Real-Time PM_{2.5} Retrievals from Multiple Data
583 Sources. *Environ Sci Technol* 2021; **55**(17): 12106–15.
- 584 38. Xue T, Zheng Y, Geng G, et al. Estimating Spatiotemporal Variation in Ambient Ozone Exposure during 2013–
585 2017 Using a Data-Fusion Model. *Environ Sci Technol* 2020; **54**(23): 14877-88.
- 586 39. Didan K, Munoz AB, Solano R, Huete A. MODIS vegetation index user's guide (MOD13 series). *Vegetation Index*
587 *and Phenology Lab* 2015: 35.
- 588 40. Wei J, Liu S, Li Z, et al. Ground-Level NO₂ Surveillance from Space Across China for High Resolution Using
589 Interpretable Spatiotemporally Weighted Artificial Intelligence. *Environ Sci Technol* 2022; **56**(14): 9988-98.
- 590 41. Wei J, Li Z, Wang J, Li C, Gupta P, Cribb M. Ground-level gaseous pollutants (NO₂, SO₂, and CO) in China: daily
591 seamless mapping and spatiotemporal variations. *Atmos Chem Phys* 2023; **23**(2): 1511-32.
- 592 42. Xiao T, Lei K, Jiang Z, et al. A Six-year long High-resolution Air Quality Reanalysis Dataset over China from 2013

- 593 to 2018 (monthly and annual version). Science Data Bank; 2021.
- 594 43. Anderson GB, Bell ML, Peng RD. Methods to calculate the heat index as an exposure metric in environmental
595 health research. *Environ Health Persp* 2013; **121**(10): 1111-9.
- 596 44. Areal AT, Zhao Q, Wigmann C, Schneider A, Schikowski T. The effect of air pollution when modified by
597 temperature on respiratory health outcomes: A systematic review and meta-analysis. *Sci Total Environ* 2022; **811**:
598 152336.
- 599 45. Hersbach H, Bell B, Berrisford P, et al. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological*
600 *Society* 2020; **146**(730): 1999-2049.
- 601 46. Malley CS, Henze DK, Kuylenstierna JCI, et al. Updated Global Estimates of Respiratory Mortality in Adults ≥ 30
602 Years of Age Attributable to Long-Term Ozone Exposure. *Environ Health Persp* 2017; **125**(8): 087021.
- 603
- 604

605 **EXTENDED DATA**

606



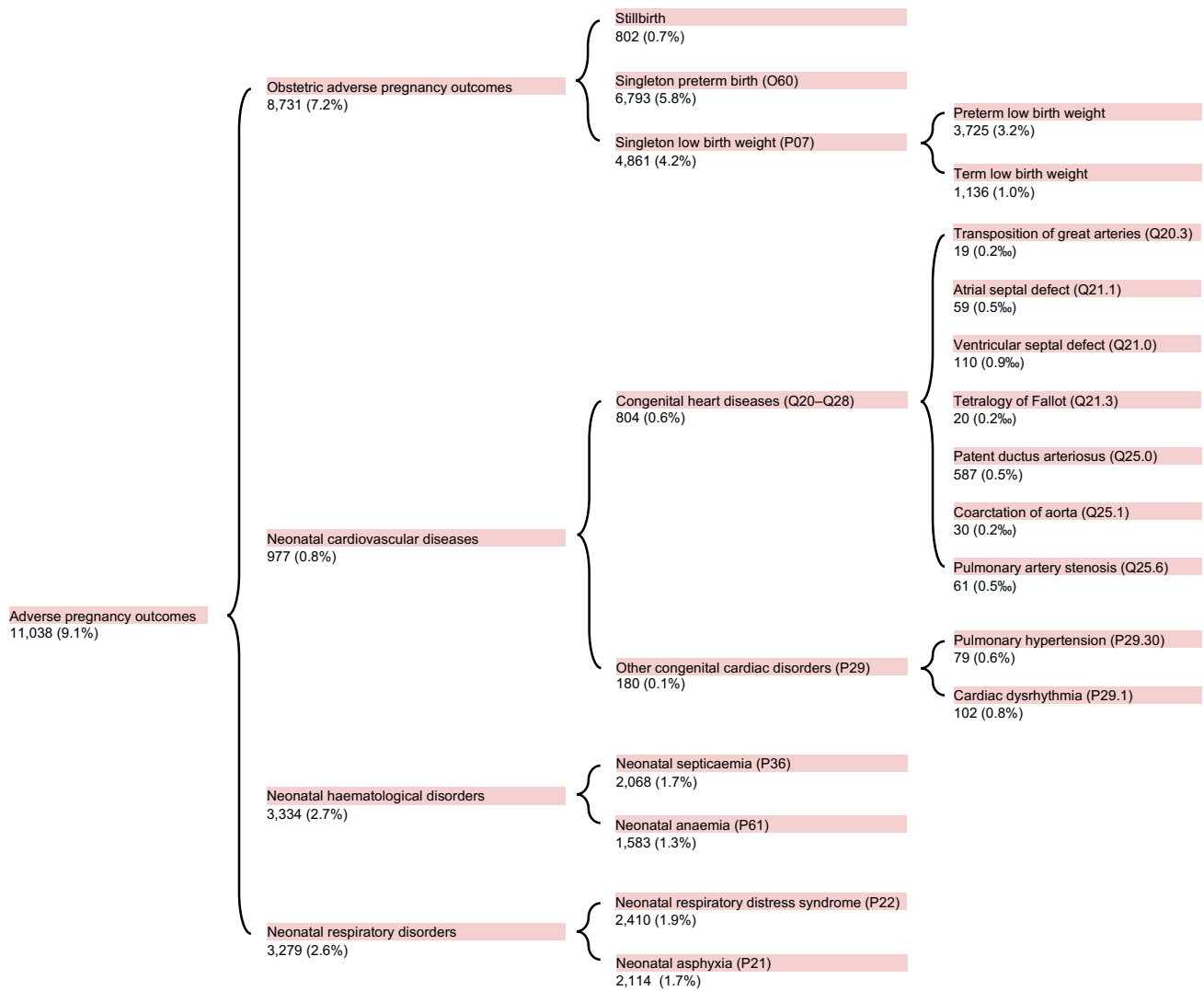
607

608 **Extended Data Fig. 1 | Incidence of adverse maternal health outcomes by hierarchical categorisation.**

609 Studied maternal diseases encompass three major categories: primary CVDs, primary haematological diseases, and pregnancy-induced
610 CVDs. The diagnosis of specific diseases corresponds to the International Classification of Diseases 10th Revision (ICD-10) coding
611 system. The number of cases and incidence rates for each hierarchical level of disease are indicated beneath the disease labels.

612

613



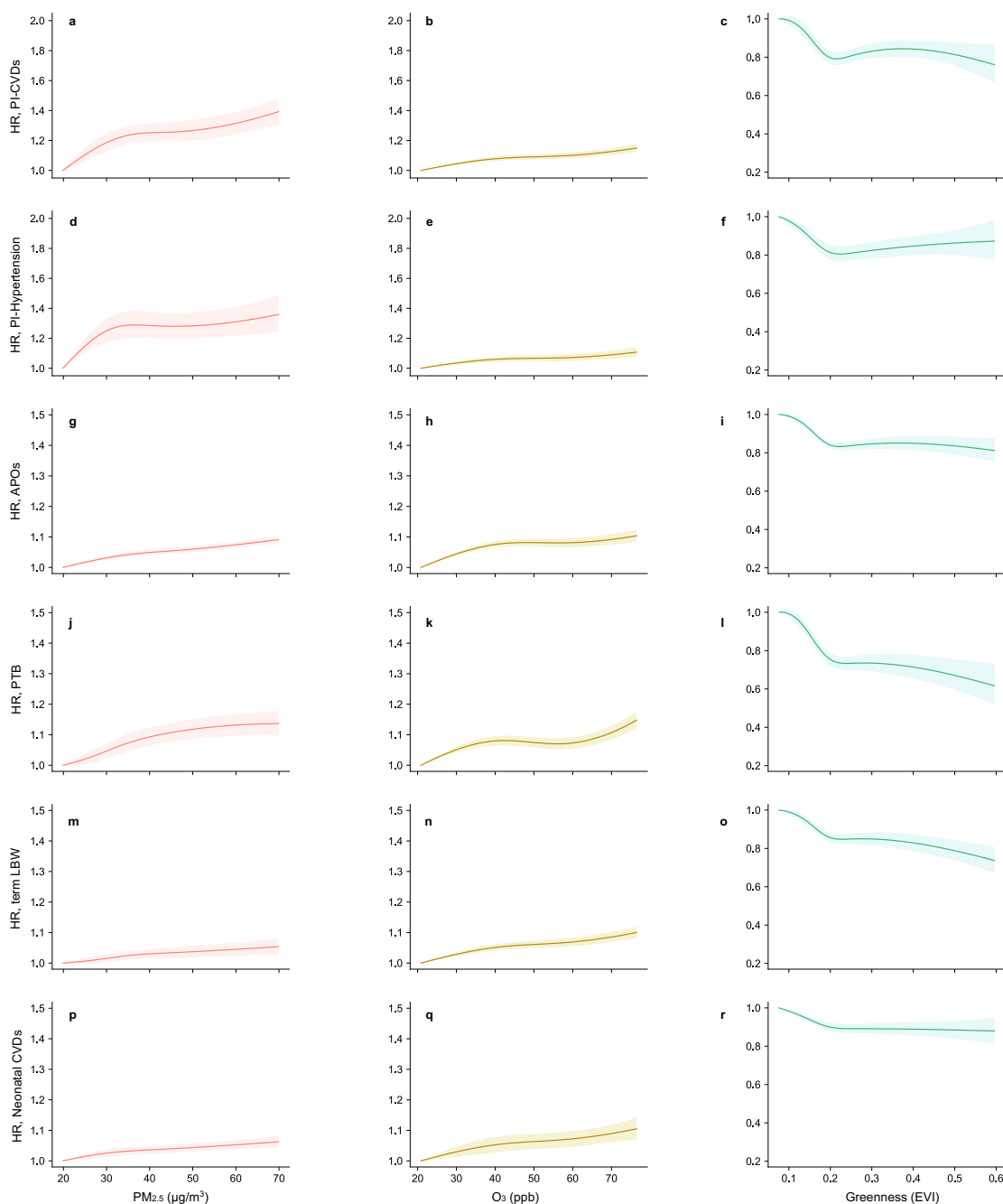
614

615 **Extended Data Fig. 2 | Incidence of adverse pregnancy outcomes by hierarchical categorisation.**

616 Adverse pregnancy outcomes encompass four aspects: obstetric anomalies, neonatal CVDs, neonatal haematological diseases, and
617 neonatal respiratory disorders. The figure configuration follows the layout of [Extended Data Fig. 1](#).

618

619



620

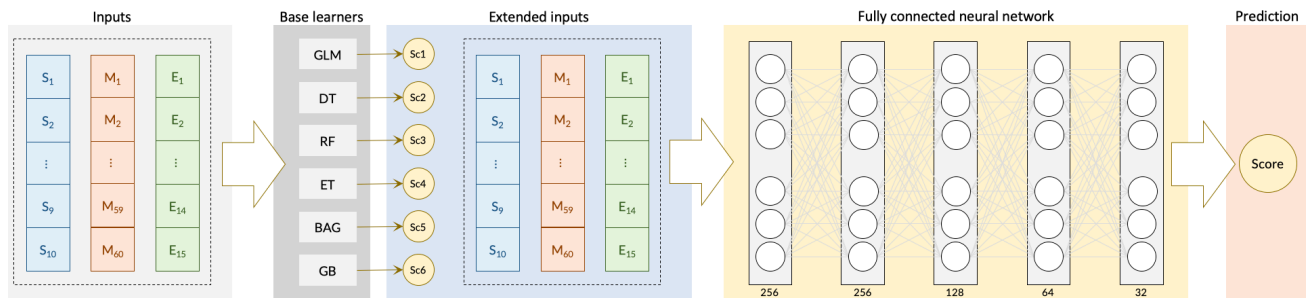
621 **Extended Data Fig. 3 | Risk association curves of environmental exposures on adverse maternal**
622 **pregnancy-induced cardiovascular complications and adverse pregnancy outcomes.**

623 Restricted cubic spline regression models with no more than 4 degrees of freedom are applied to investigate the curved risk association
624 trends (quantified in hazard ratio, HR) between maternal PM_{2.5}, O₃, greenness exposure (presented in columns) and adverse health
625 outcomes. **a-c**, pregnancy-induced cardiovascular diseases; **d-f**, pregnancy-induced hypertension; **g-i**, adverse pregnancy outcomes, **j-l**,
626 preterm birth, **m-o**, term low birth weight, **p-r**, neonatal cardiovascular diseases. Thresholds are defined as the lowest 5th percentiles of
627 the exposure levels among the studied cohort participants, as 19.8 µg/m³, 20.9 ppb, and 0.076 EVI for PM_{2.5}, O₃, and greenness,
628 respectively. The trend observation ranges are determined by 1st-99th percentiles, as 16.4–76.4 µg/m³, 16.1–76.2 ppb, and 0.044–0.542
629 EVI for PM_{2.5}, O₃, and greenness, respectively.

630

631

632



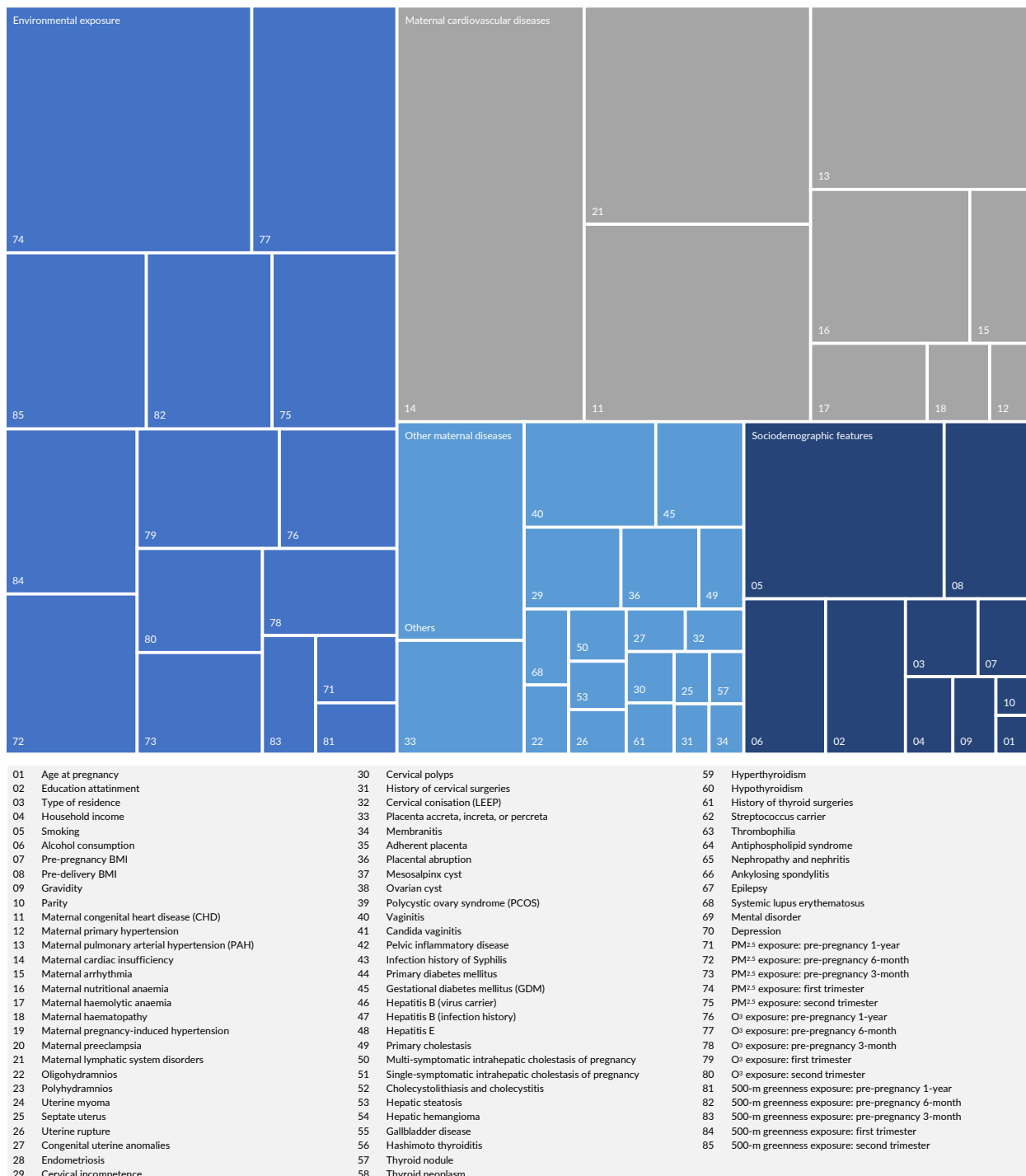
633

634 **Extended Data Fig. 4 | Schematic diagram of ensemble-learning-based risk prediction algorithm framework for adverse pregnancy outcomes.**

635 The prediction targets of the algorithm are i) total obstetric anomalies and ii) total neonatal CVDs, without distinguishing any subsets.
636 Initially, the algorithm takes 10 socioeconomic features (denoted as S), 60 medical diagnostic records (M), and 15 environmental
637 exposure factors (E) as input layers. It conducts first-stage risk prediction using six base learners: generalised linear model (GLM),
638 decision tree classifier (DT), random forest classifier (RF), extra-tree classifier (ET), bootstrap aggregating classifier (BAG), and
639 gradient boosting classifier (GB). The predicted risk scores (Sc) obtained in the first-stage algorithm, together with the raw initial
640 inputs, serve as the new input layers for the second-stage fully connected neural network (FCNN). The FCNN comprises five hidden
641 layers, each consisting of 256, 256, 128, 64, and 32 nodes, respectively. The outputs of the second-stage algorithm ultimately represent
642 the predicted occurrence probabilities for the two designed prediction targets.

643

644



645

646

Extended Data Fig. 5 | Treemap chart of weighted risk scores of obstetric adverse pregnancy outcomes.

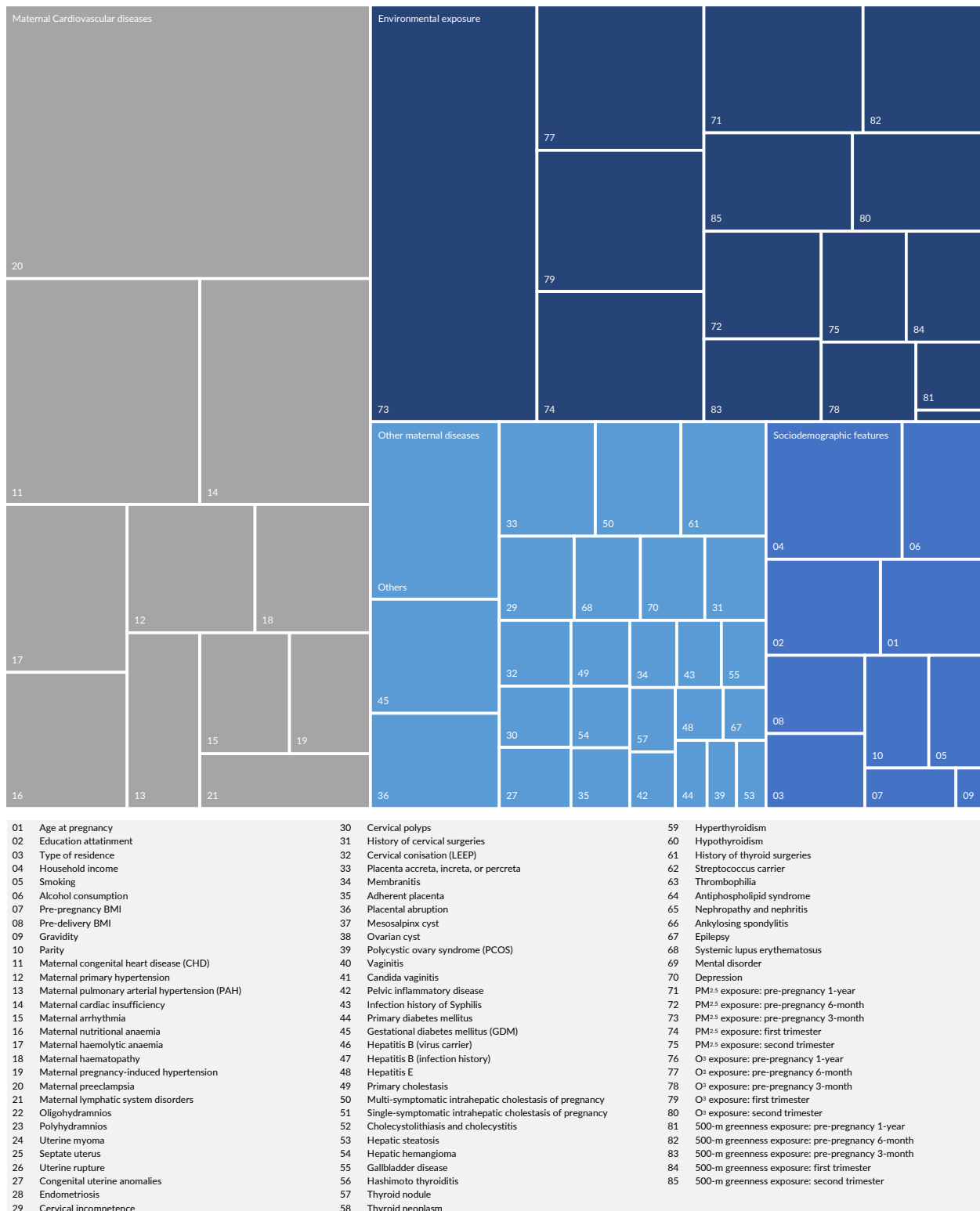
647

The treemap illustrates the weights of the 85 risk factors used for prediction via the self-assessment form. These risk factors are categorised into four groups: sociodemographic features, environmental exposure, maternal CVDs, and other maternal diseases, comprising 12.6%, 38.0%, 34.5%, and 14.9% of the total weight, respectively. In clinical practice, pregnant women who develop pregnancy-induced hypertension or even preeclampsia are almost inevitably subjected to caesarean section, leading to medically indicated preterm birth. Therefore, maternal pregnancy-induced hypertension and preeclampsia were excluded in risk weighting.

651

652

653



654

655 **Extended Data Fig. 6 | Treemap chart of weighted risk scores of neonatal cardiovascular diseases.**

656 The layout of the figure follows the configuration presented in Extended Data Fig. 5. The four groups of risk factors: sociodemographic
657 features, environmental exposure, maternal CVDs, and other maternal diseases, account for weights of 10.7%, 32.5%, 37.3%, and 19.5%,
658 respectively.

659

660

661 **Extended Data Table 1 | Statistical summary of ZEBRA maternity cohort profile, 2013–2022.**

Characteristics		Mean (SD) or Cases (%)
Maternal age at delivery		31.2 (4.3)
Height (cm)		160.6 (4.9)
Weight before pregnancy (Kg)		53.4 (8.0)
BMI before pregnancy		21.1 (2.8)
Weight before delivery (Kg)		67.9 (8.4)
BMI before delivery		26.3 (3.0)
Gestational age (day)		272.6 (13.3)
Menarche age		13.9 (1.4)
Gravidity (including current pregnancy)	1	50,089 (41.1%)
	2	36,255 (29.7%)
	3–5	33,684 (27.6%)
	>5	1,864 (1.5%)
Parity (including current pregnancy)	1	72,412 (59.4%)
	2	46,667 (38.3%)
	3	2,653 (2.2%)
	>3	160 (0.1%)
Ethnicity	Hàn	120,170 (98.6%)
	Manchu	246 (0.2%)
	Tǔ-Jiā	240 (0.2%)
	Huí	224 (0.2%)
	Zhuàng	187 (0.2%)
	Shē	181 (0.1%)
	Others	644 (0.5%)
Occupation	Office staff	97,392 (79.9%)
	Worker	14,505 (11.9%)
	Professional and specialist	3,291 (2.7%)
	Farmer	975 (0.8%)
	Medical staff	366 (0.3%)
	Student	123 (0.1%)
	Others	5,240 (4.3%)
Education attainment	High school or below	19,511 (16.0%)
	College education	23,249 (19.1%)
	Undergraduate	62,083 (50.9%)
	Graduate or above	17,049 (14.0%)
Type of residence	Urban	93,846 (77.0%)
	Rural	28,046 (23.0%)
Type of medical insurance	NCMS	8,644 (7.1%)
	URBMI	7,530 (6.2%)
	UEBMI	105,718 (86.7%)
Alcohol and smoking history	Alcohol	124 (0.1%)
	Smoking	35 (0.3%)
	Second-hand smoking	2,214 (1.8%)

662

663

664 **Extended Data Table 2 | Statistical summary of medical history.**

History of diseases and medical treatments	Cases (%)
Uterine fibroids	7,951 (6.5%)
Septate uterus	613 (0.5%)
Uterine rupture	474 (0.4%)
Congenital uterine malformation	293 (0.2%)
Cervical incompetence	1,107 (0.9%)
History of cervical surgery	687 (0.6%)
Cervical polyps	427 (0.4%)
Cervical conisation (LEEP)	314 (0.3%)
Membranitis	527 (0.4%)
Placenta accreta, increta, or percreta	825 (0.7%)
Morbidly adherent placenta	3,369 (2.8%)
Placental abruption	2,675 (2.2%)
Mesosalpinx cyst	2,758 (2.3%)
Ovarian cyst	2,136 (1.8%)
Polycystic ovary syndrome (PCOS)	356 (0.3%)
Vaginitis	552 (0.5%)
Pelvic inflammatory disease (PID)	455 (0.4%)
Systemic lupus erythematosus	71 (0.1%)
Syphilis (infection history)	423 (0.3%)
Gestational diabetes mellitus	290 (0.2%)
Antiphospholipid syndrome (APS)	902 (0.7%)
Hepatitis B (virus carrier)	6,825 (5.6%)
Hepatitis B (infection history)	5,349 (4.4%)
Hepatic haemangioma	64 (0.1%)
Gallbladder diseases	386 (0.3%)
Single-symptomatic intrahepatic cholestasis of pregnancy	2,864 (2.3%)
Multi-symptomatic intrahepatic cholestasis of pregnancy	805 (0.7%)
Cholecystolithiasis and cholecystitis	313 (0.3%)
Hepatic steatosis (HS)	61 (0.1%)
Other liver diseases or anomalies	1,932 (1.6%)
Hashimoto thyroiditis	1,043 (0.9%)
Thyroid cancer	720 (0.6%)
Hyperthyroidism	231 (0.2%)
Hypothyroidism	406 (0.3%)
History of thyroid surgery	402 (0.3%)
Kidney diseases	255 (0.2%)
Endometriosis	815 (0.7%)
Thrombophilia	2,031 (1.7%)
Ankylosing spondylitis	160 (0.1%)
Epilepsy	88 (0.1%)
Mental disorders	82 (0.1%)
Depression	187 (0.2%)

665

666

667 **Extended Data Table 3 | Sensitivity analyses on risk associations by adjusting for combinations of**
668 **confounders.**

Confounder Adjustment	PM _{2.5} (10 µg/m ³)	O ₃ (10 ppb)	Greenness (0.1 EVI)
Maternal pregnancy-induced cardiovascular diseases			
Crude HR	1.053 (1.046–1.060)	1.006 (1.002–1.011)	0.946 (0.931–0.962)
Combo 1: Physiological and lifestyle characteristics, including age at pregnancy, BMI, ethnicity, alcohol and smoking history.	1.050 (1.043–1.058)	1.008 (1.003–1.013)	0.947 (0.931–0.963)
Combo 2: Combo 1 + socioeconomic status, including urban/rural residence, medical insurance, occupation, education attainment.	1.054 (1.047–1.062)	1.008 (1.003–1.013)	0.929 (0.914–0.945)
Combo 3: Combo 2 + obstetric and relevant surgical history, including gravidity, parity, history of cervical, hysteroscopic and placental surgeries.	1.057 (1.050–1.064)	1.008 (1.003–1.013)	0.930 (0.915–0.946)
Combo 4: Combo 3 + non-cardiovascular medical history, including liver and gallbladder diseases, thyroid diseases, metabolic disorders, immune system diseases, mental disorders, etc.	1.057 (1.050–1.064)	1.008 (1.003–1.013)	0.931 (0.916–0.947)
Combo 5: Combo 4 + cardiovascular and haematological medical history.	1.055 (1.048–1.062)	1.009 (1.005–1.014)	0.931 (0.916–0.947)
Combo 6: Combo 5 + multiple studied environmental exposures, as PM _{2.5} , O ₃ , and greenness.	1.070 (1.061–1.079)	1.029 (1.024–1.035)	0.968 (0.951–0.985)
Combo 7: Combo 6 + additional environmental exposures, as NO ₂ , SO ₂ , CO, temperature and humidity.	1.073 (1.064–1.082)	1.027 (1.022–1.033)	0.964 (0.948–0.982)
Adverse pregnancy outcomes			
Crude HR	1.013 (1.010–1.016)	1.010 (1.008–1.012)	0.974 (0.968–0.980)
Combo 1: Physiological and lifestyle characteristics, including age at pregnancy, BMI, ethnicity, alcohol and smoking history.	1.012 (1.009–1.015)	1.010 (1.009–1.012)	0.977 (0.971–0.983)
Combo 2: Combo 1 + socioeconomic status, including urban/rural residence, medical insurance, occupation, education attainment.	1.015 (1.013–1.018)	1.011 (1.009–1.013)	0.955 (0.949–0.961)
Combo 3: Combo 2 + obstetric and relevant surgical history, including gravidity, parity, history of cervical, hysteroscopic and placental surgeries.	1.018 (1.015–1.021)	1.010 (1.008–1.012)	0.951 (0.945–0.957)
Combo 4: Combo 3 + non-cardiovascular medical history, including liver and gallbladder diseases, thyroid diseases, metabolic disorders, immune system diseases, mental disorders, etc.	1.018 (1.015–1.021)	1.010 (1.008–1.012)	0.952 (0.946–0.958)
Combo 5: Combo 4 + cardiovascular and haematological medical history.	1.016 (1.013–1.018)	1.011 (1.009–1.013)	0.951 (0.945–0.957)
Combo 6: Combo 5 + multiple studied environmental exposures, as PM _{2.5} , O ₃ , and greenness.	1.019 (1.016–1.023)	1.016 (1.014–1.019)	0.960 (0.954–0.967)
Combo 7: Combo 6 + additional environmental exposures, as NO ₂ , SO ₂ , CO, temperature and humidity.	1.015 (1.012–1.019)	1.017 (1.015–1.019)	0.961 (0.954–0.967)

669

670

671 **Extended Data Table 4 | Direct and indirect effects of maternal pregnancy-induced cardiovascular**
672 **diseases as mediator between environmental exposure and adverse pregnancy outcomes.**

673 Total adverse pregnancy outcomes (APO) together with three subsets, preterm birth (PTB), low birth weight (LBW), and neonatal
674 cardiovascular diseases (CVD) are assessed for mediation effects via maternal pregnancy-induced CVDs. Environmental exposures
675 include PM_{2.5}, O₃ and greenness, quantified in 10-μg/m³, 10-ppb, and 0.1-EVI increment, respectively. Effects are quantified in β, log-
676 transformed hazard ratio (HR), estimated from Cox proportional hazard regression models with 95% confidence intervals (CI).
677 Proportions of direct and indirect effects are estimated in percentages (%).

		PM _{2.5} (10 μg/m ³)			O ₃ (10 ppb)			Greenness (0.1 EVI)		
		β	95% CI	%	β	95% CI	%	β	95% CI	%
APO	total effect	0.026	(0.022–0.029)		0.022	(0.020–0.024)		-0.048	(-0.054 to -0.041)	
	direct effect	0.014	(0.011–0.018)	55.6	0.015	(0.013–0.017)	65.6	-0.037	(-0.043 to -0.030)	76.9
	indirect effect	0.011		44.4	0.008		34.4	-0.011		23.1
PTB	total effect	0.013	(0.008–0.017)		0.022	(0.020–0.025)		-0.029	(-0.038 to -0.020)	
	direct effect	0.006	(0.002–0.011)	49.1	0.012	(0.010–0.015)	55.2	-0.022	(-0.031 to -0.013)	75.1
	indirect effect	0.006		50.9	0.010		44.8	-0.007		24.9
LBW	total effect	0.039	(0.035–0.042)		0.031	(0.029–0.034)		-0.077	(-0.086 to -0.069)	
	direct effect	0.023	(0.019–0.027)	60.9	0.020	(0.018–0.023)	65.5	-0.068	(-0.077 to -0.059)	87.9
	indirect effect	0.015		39.1	0.011		34.5	-0.009		12.1
CVD	total effect	0.062	(0.051–0.073)		0.043	(0.036–0.049)		-0.027	(-0.049 to -0.005)	
	direct effect	0.048	(0.038–0.059)	78.2	0.036	(0.029–0.042)	84.3	-0.024	(-0.046 to -0.002)	86.7
	indirect effect	0.014		21.8	0.007		15.7	-0.004		13.3

678

679 **Data and code availability**

680 The information of the ZEBRA maternity cohort participants for privacy protection it is not disclosed for public use.
681 Researchers interested in accessing the data are encouraged to contact the principal investigators with a brief research
682 proposal. Access to the data will be granted after approval by the ZEBRA committee and the Health Commission of
683 Zhejiang Province. The codes for analysis can be shared upon request.

684 **Acknowledgements**

685 This study is funded by the Zhejiang Province Health Innovative Talent Project (A0466), International Cooperation
686 Seed Program of Women's Hospital, Zhejiang University (GH2022B008-01), Key Projects of the Science and Technology
687 Co-construction by National Administration of Traditional Chinese Medicine and Zhejiang Provincial Administration of
688 Traditional Chinese Medicine (GZY-ZJ-KJ-23082), Australian Research Council (DP210102076), and Australian National
689 Health and Medical Research Council (APP2000581). Jason Sun also thanks generous support from the U.S. Fulbright
690 Program. Yuming Guo is supported by a Career Development Fellowship of the Australian National Health and Medical
691 Research Council (APP1163693). Special appreciations to Professor Weiguo Lu, the director of the Key Laboratory of
692 Women's Reproductive Health, for his comprehensive supports in establishing and coordinating the maternity cohort.

693 **Author contributions**

694 J.S., Y.G., and X.B. conceptualised and designed the study. W.X. and X.B. were responsible for the administration of
695 the maternity cohort. H.T., H.Z., Q.X., Y.T., and other non-academic members from the ZEBRA collaborative group
696 collected, censored, and pre-processed medical records and associated data for analysis. J.S. performed analyses and wrote
697 the manuscript with comprehensive supports from all authors. K.R.D., K.T., E.X.L., and L.P.S. supervised the study design,
698 reviewed and edited the manuscript.

699 **Competing interests**

700 The authors declare that they have no known competing financial interests or personal relationships that could have
701 appeared to influence the work reported in this paper.

702 **Additional information**

703 Supplementary Materials include 6 long-list supplementary contents and 4 tables in 55 pages.

704

705

706 **SUPPLEMENTARY TABLES**

707

708 [Supplementary Table S1](#) | Self-directed risk assessment questionnaire for early-stage forecasting of obstetric
709 adverse pregnancy outcomes and neonatal congenital cardiovascular diseases.

710 [Supplementary Table S2](#) | Scoring for risk factors of self-directed questionnaire for early-stage risk forecasting
711 of obstetric adverse pregnancy outcomes and neonatal congenital cardiovascular diseases.

712 [Supplementary Table S3](#) | A sample filled self-directed questionnaire with the highest risk score of obstetric
713 ad-verse pregnancy outcomes.

714 [Supplementary Table S4](#) | A sample filled self-directed questionnaire with the highest risk score of neonatal
715 cardiovascular diseases.

716 [Supplementary Table S5](#) | Sensitivity analysis: Fully adjusted risk associations between environmental
717 exposure and adverse health outcomes by bootstrap resampling.

718 [Supplementary Table S6](#) | Sensitivity analysis: Fully adjusted risk associations between maternal
719 cardiovascular diseases (primary and pregnancy-induced) and adverse pregnancy outcomes (obstetric and
720 neonatal diseases) by bootstrap resampling.

721 [Supplementary Table S7](#) | Sensitivity analysis: Risk associations between environmental exposure and adverse
722 health outcomes by different combination of multi-factor adjustment.

723 [Supplementary Table S8](#) | Sensitivity analysis: Robustness of estimated risk associations at Zhejiang
724 Provincial scale and China nationwide scale.

725 [Supplementary Table S9](#) | STROBE checklist: Checklist of items that should be included in reports of
726 observational studies.

727 [Supplementary Table S10](#) | ZEBRA Collaborative Group full roster.

728