

## Supplementary material

### **Definition of the metabolic syndrome**

The MS was defined using NCEP-ATP III (National Cholesterol Education Program Adult Treatment Panel III) clinical criteria except for including: abdominal obesity (increased waist circumference: >102 cm for men; >88 cm for women), raised triglycerides ( $\geq 1.50\text{g/L}$  ( $\geq 1.69\text{ mmol/L}$ ), reduced HDL cholesterol ( $< 40\text{ mg/dL}$  for (men) ( $< 1.04\text{ mmol/L}$ ),  $< 50\text{ mg/dL}$  for women ( $< 1.29\text{ mmol/L}$ ), elevated blood pressure (BP  $\geq 130$  and/or  $\geq 85\text{ mm Hg}$ ), and raised plasma glucose. If three out of the five listed characteristics were present, a diagnosis of metabolic syndrome was made. In our analyses, BP was calculated as the average of three readings. Waist circumference was considered only if the measurement was noted by the cohort team as being “reliable”. Plasma glucose which was not recorded was replaced by Glycated haemoglobin (HbA1c)  $\geq 6.5\%$ .

### **Statistical analyses**

The imputation model included the exposure variables and covariates used in the logistic regression models, as well as other correlated variables which were likely to improve the imputation model but that were not used in analyses: marital status of the mother, socio-economic group of the mother’s father, multiple pregnancy. The outcome variable was omitted from the imputation model. Ten dataset imputations were run. The logistic regression models were carried out on the imputed data in order to decrease the potential bias of our estimated ORs and increase analytical power. Standard errors were calculated using Rubin’s rules (1). P-values  $\leq 0.05$  were considered as statistically significant. Maternal age of birth was treated as a continuous variable, as a fractional polynomial model showed there was no evidence that the change in incidence over time was not linear (2).

### **Analyses of missing data**

Having missing data or not for the covariates was analysed in terms of mode of delivery and in terms of the outcome measure. Missing data for mother’s BMI before pregnancy was associated with being born by planned caesarean (10.8% vs 8.7% and 5.4%  $p=0.03$ ); Missing data for toxemia was associated with being born by unplanned caesarean (10.6% vs 6.0% and 4.8%  $p=0.029$ ). Missing data for birthweight was associated with unplanned caesarean birth (20.19% versus 16.9% and 12.8%  $p=0.05$ ) and with having the MS (14.4% vs 12.5%  $p=0.029$ ). Missing data for PROM was associated with a vaginal birth (5.9% versus 4.8% and

0%  $p=0.035$ ) and with having the MS (6.6% s 5.4  $p=0.049$ ). No clear pattern emerges from these analyses of missing data that suggests a systematic bias, except perhaps for persons with data missing for birthweight which was associated with both unplanned caesarean birth and the metabolic syndrome. However, the analyses using multiple imputations would adjust for this possible bias.

1. Rubin D. Multiple imputation for nonresponse in surveys. New York: Wiley, 1987 j'ai trouvé ça sur pubmed: Mehrotra DV, Li X, Liu J, Lu K. Analysis of longitudinal clinical trials with missing data using multiple imputation in conjunction with robust regression. *Biometrics* 2012 Dec;68(4):1250-9.
2. Royston P, Altman DG. Using fractional polynomials to model curved relationships, *Stata Technical Bulletin* 1994; 21:11-24