

Supplementary file

Table S1: Diagnosis codes used to identify primary outcome

Outcome	Content	Data extracted from CPC/MBR/DNPR/CPR Diagnosis codes used for extraction (ICD-10 codes) (SKS-codes)
Primary	Occurrence of PPRM	MBR: pprom, PrimærVandafgang (rupture of fetal membranes indicated as either PROM or PPRM), DO420, DO422, DO424. DNPR: DO420, DO422, DO424.

CPC: Copenhagen Pregnancy Cohort, MBR: Medical Birth Registry, DNPR: Danish National Patient Registry,
CPR: The Danish Civil Registration System.

Table S2: Socio-demographic characteristics of the women in the study population

Maternal characteristics	Multiparous women n= 14,198 (39.1%)
Maternal age, median years (IQR) <i>Missing</i>	33.5 (30.9-36.3) 13
Pre-pregnancy BMI, median kg/m ² (IQR) ¹ <i>Missing</i>	21.9 (20.2-24.3) 587
Pre-pregnancy BMI, kg/m ² (IQR) <18.5 18.5-24.9 25-29.9 ≥30 <i>Missing</i>	750 (5.3) 10,112 (71.2) 1,999 (14.1) 750 (5.3) 587
Education level, n (%) Higher degree (≥ 5 years) Intermediate degree (3-4 years) Short degree (1-2 years) Technical degree Compulsory education <i>Missing</i>	7,633 (53.8) 4,078 (28.7) 843 (5.9) 426 (3.0) 758 (5.3) 460

Assisted reproductive technology, n (%) None Egg donation IVF/ICSI Hormone stimulation <i>Missing</i>	13,381 (94.2) 9 (0.1) 140 (1.0) 668 (4.7) 0
Smoking during the first trimester, n (%) Yes No <i>Missing</i>	193 (1.4) 13,742 (96.8) 263
Physical activity in the first trimester, n (%) Yes No <i>Missing</i>	6,934 (48.8) 7,264 (51.2) 0
Continent of origin, n (%) Africa South and East Asia North America South America Oceania Europe <i>Missing</i>	218 (1.5) 877 (6.2) 45 (0.3) 94 (0.7) 23 (0.2) 12,941 (91.1) 0
Cohabitation, n (%) Yes No <i>Missing</i>	13,321 (93.8) 757 (5.3) 120
Chronic medical disease, n (%) None Somatic Psychiatric Somatic and psychiatric <i>Missing</i>	10,735 (75.6) 3,075 (21.7) 270 (1.9) 118 (0.8) 0
WHO-5 Well-being index score, median points (IQR) <i>Missing</i>	64 (52-72) 196
Previous cervical conization, n (%) Yes No <i>Missing</i>	264 (1.9) 13,934 (98.1) 0
Previous abortion, n (%) Yes No <i>Missing</i>	3,869 (27.3) 10,329 (72.7) 0
Previous sPTB, n (%) Yes No <i>Missing</i>	895 (6.3) 13,303 (93.7) 0

Previous PPROM, n (%) Yes No <i>Missing</i>	429 (3) 13,283 (93.6) 486
Previous bleeding in early pregnancy, n (%) Yes No <i>Missing</i>	63 (0.4) 14,135 (99.6) 0

BMI: Body Mass Index, sPTB: Spontaneous preterm birth, PPROM: Preterm prelabour rupture of membranes

Table S3: Specifications of predictor variables included in the SuperLearner model

Predictor variable	Content	Type of variable	Included for nulliparous women/multiparous women	Data extracted from CPC/MBR/DNPR/CPR Diagnosis codes used (ICD-10 codes) (SKS-codes) and variable names used for extraction
<i>Risk factors</i>				
Maternal age	Age when filling out the questionnaire	Continuous	Both	CPC MBR: alder_moder, alder_mor
Pre-pregnancy BMI	Pre-pregnancy BMI (kg/m ²) Categorized by pre-pregnancy BMI: <18.5 kg/m ² 18.5-24.9 kg/m ² 25-29.9 kg/m ² ≥30 kg/m ²	Continuous Categorical	Both Both	CPC MBR: bmi_moder, bmi_mor
Educational level	Categorized by length of education (years): Higher degree (≥ 5 years) Intermediate degree (3-4 years) Short degree (1-2 years) Technical degree Compulsory education	Categorical	Both	CPC

Assisted reproductive technology	Categorized by treatment and ranking: 1. Egg donation 2. IVF/ICSI 3. Hormone treatment	Categorical	Both	CPC
Smoking in the first trimester	Smoking in the first trimester (yes/no)	Dichotomous	Both	CPC
Ethnicity	Categorized by continent of origin for immigrants and descendants of immigrants Africa Asia North America South America Oceania Europe	Categorical	Both	CPC CPR: oprindelses_land
Cohabitation	Cohabitations status when filling out questionnaire	Dichotomous	Both	CPC
Chronic medical disease	Categorized by origin of disease: None Somatic Psychiatric Somatic and psychiatric	Categorical	Both	CPC
Previous conization	Previous conization prior to pregnancy (yes/no)	Dichotomous	Both	CPC MBR: DO344B DNPR: KLD03

Previous spontaneous preterm birth	Previous obstetric history of preterm birth (yes/no)	Dichotomous	Multiparous	CPC MBR: DZ358B
Previous PPROM	Previous obstetric history of PPROM	Dichotomous	Multiparous	MBR: pprom, PrimærVandaf- ganga (indicated as either PROM or PPROM), DO420, DO422, DO424. DNPR: DO420, DO422, DO424.
Bleeding in early pregnancy	Episode/episodes of bleeding in early pregnancy (un- specified GW) (yes/no)	Dichotomous	Multiparous	MBR: DO20
<i>Potential risk factors</i>				
Physical activity in the first tri- mester	Physically active in the first trimester (yes/no)	Dichotomous	Both	CPC
WHO-5 Well-being Index scale	Assessment of well-being index using a point system. The higher the score the lower risk of stress and de- pression.	Continuous	Both	CPC
Previous abortion	Previous abortion prior to pregnancy (yes/no)	Dichotomous	Both	CPC

CPC: Copenhagen Pregnancy Cohort, MBR: Medical Birth Registry, DNPR: Danish National Patient Registry, CPR: The Danish Civil Registration System, BMI: Body Mass Index,

PPROM: Preterm prelabour rupture of membranes, sPTB: Spontaneous preterm birth, GW: Gestational Weeks, WHO: World Health Organization.

Table S4: Overview of incidences of outcomes			
Incidence of outcomes	Nulliparous women n= 22,071 (60.9)	Multiparous women n= 14,198 (39.1)	Study population (n=36,269)
PPROM, n (%)	1218 (5.5)	329 (2.3)	1547 (4.3)
sPTB <37 gestational weeks, n (%)	1,062 (4.8)	556 (3.9)	1,618 (4.5)
sPTB <34 gestational weeks, n (%)	343 (1.6)	147 (1)	490 (1.3)

PPROM: preterm prelabour rupture of membranes, sPTB: spontaneous preterm birth.

Table S5 Brier scores of candidate models for nulliparous women, highlighted by the most well-performing model in bold writing

	Brier score										
Outcome/ Model	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	Null model
PPROM	0.0524378	0.0524378	0.0524378	0.0524378	0.0524483	0.0524483	0.0524167	0.0524167	0.0524746	0.0524746	0.0525464
sPTB <37 gestational weeks	0.0452229	0.0452254	0.0452226	0.0452233	0.0451988	0.0452013	0.0452719	0.0452577	0.0452042	0.0452084	0.0461796
sPTB <34 gestational weeks	0.0143371	0.014338	0.0143367	0.0143375	0.0143342	0.0143378	0.0143421	0.0143408	0.0143345	0.0143378	0.0143390

PPROM: Preterm prelabour rupture of fetal membranes, sPTB: Spontaneous preterm birth.

Table S6: Brier scores of candidate models for multiparous women, highlighted by the most well-performing model in bold writing

	Brier score											
Out-come/ Model	M0	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	Null model
PPROM	0.0223140	0.022325	0.022325	0.0223467	0.0223492	0.0223394	0.0223393	0.0223415	0.022341	0.0233140	0.0223228	0.0223399
sPTB <37 ge- stational weeks	0.0345890	0.03457136	0.0345773	0.0345707	0.0345689	0.0345701	0.0345725	0.0347841	0.0345480	0.0345672	0.03443419	0.0360133
sPTB <34 ge- stational weeks	0.0092621	0.00912176	0.00912141	0.0091212	0.0091214	0.0090931	0.0090931	0.0092621	0.0092036	0.0096634	0.00909699	0.00932543

PPROM: Preterm prelabour rupture of membranes, sPTB: Spontaneous preterm birth, BMI: Body Mass Index,

Table S7: Receiver operating characteristic (ROC) curve with area under the characteristic curve (AUC) calculated for each outcome		
Multiparous women		
PPROM	0.559	0.437-0.522
sPTB <37 GW	0.643	0.603-0.699
sPTB <34 GW	0.652	0.520-0.731

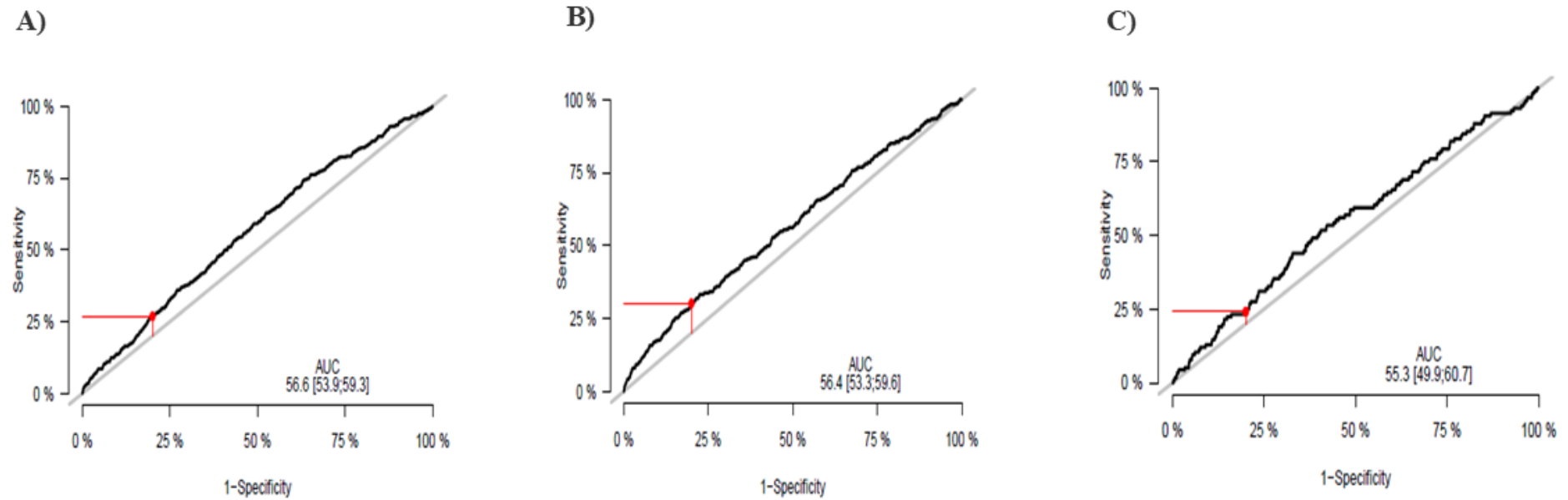
AUC: Area under the characteristic curve, CI: Confidence interval, PPROM: Preterm prelabour rupture of fetal membranes, sPTB: Spontaneous preterm birth

Figure legends:

Figure S1: Graphical representation of receiver operating characteristic (ROC) curve with area under the characteristic curve (AUC) displayed for each outcome for nulliparous women.

Figure S2: Graphical representation of receiver operating characteristic (ROC) curve with area under the characteristic curve (AUC) displayed for each outcome for multiparous women.

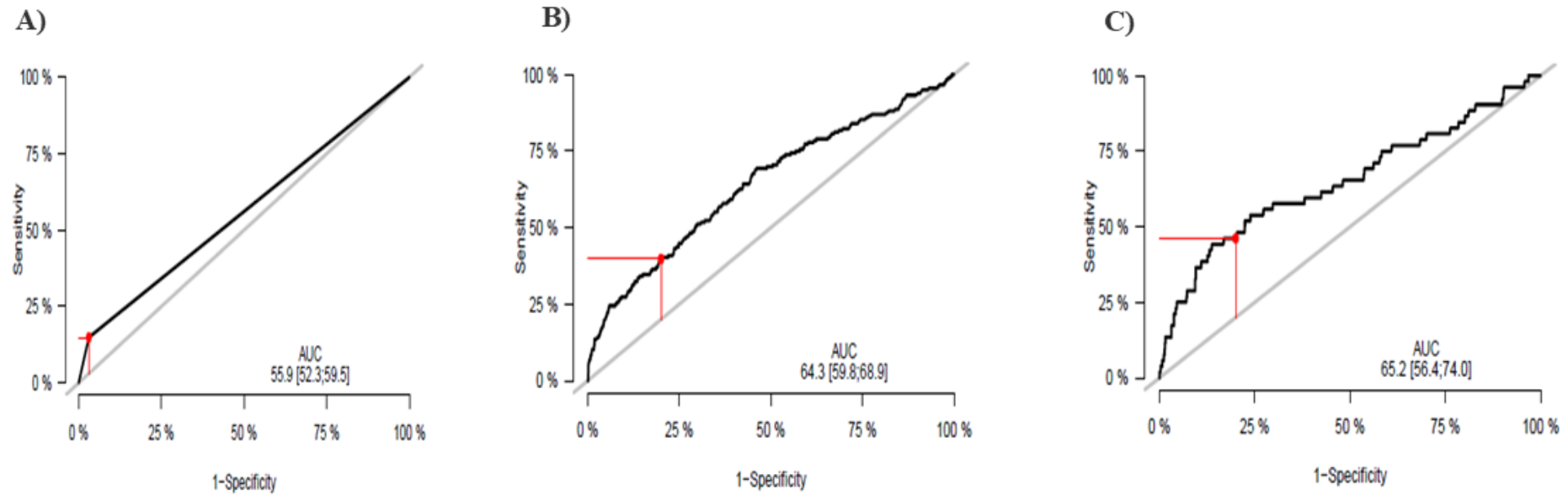
Figure S1



AUC: Area under the characteristic curve.

Graphical representation of receiver operating characteristic (ROC) curve with area under the characteristic curve (AUC) calculated for: **A)** Preterm prelabour rupture of membranes (PPROM) for nulliparous women, **B)** Spontaneous preterm birth <37 gestational weeks for nulliparous women, **C)** Spontaneous preterm birth <34 gestational weeks for nulliparous women.

Figure S2



AUC: Area under the characteristic curve.

Graphical representation of receiver operating characteristic (ROC) curve with area under the characteristic curve (AUC) calculated for: **A)** Preterm prelabour rupture of membranes (PPROM) for multiparous women,

B) Spontaneous preterm birth <37 gestational weeks for multiparous women, **C)** Spontaneous preterm birth <34 gestational weeks for multiparous women.