Received: 10 September 2018 | Revised: 10 April 2019 | Accepted: 2 May 2019 | First published online: 21 May 2019

DOI: 10.1002/ijgo.12839

CLINICAL ARTICLE

Obstetrics

WILEY International Journal of Gynecology & Obstetrics FIGO

Cluster analysis identifying clinical phenotypes of preterm birth and related maternal and neonatal outcomes from the Brazilian Multicentre Study on Preterm Birth

Renato T. Souza[1] | Jose G. Cecatti[1,*] | Renato Passini Jr[1] | Rodolfo C. Pacagnella[1] | Paulo F. Oliveira[2] | Cleide M. Silva[2] | the Brazilian Multicentre Study on Preterm Birth study group[a]

[1]Department of Obstetrics and Gynaecology, University of Campinas (UNICAMP) School of Medicine, Campinas, SP, Brazil

[2]Unit of Statistics, School of Medical Sciences, University of Campinas (UNICAMP), Campinas, Brazil

*Correspondence
Jose G. Cecatti, Department of Obstetrics and Gynecology, University of Campinas, Campinas, SP, Brazil.
Email: cecatti@unicamp.br

[a]Membership of the Brazilian Multicentre Study on Preterm Birth study group is listed at the end of the paper.

Funding Information
CNPq; FAPESP; Bill and Melinda Gates Foundation

Abstract

Objective: To explore a conceptual framework of clinical conditions associated with preterm birth (PTB) by cluster analysis, assessing determinants for different PTB subtypes and related maternal and neonatal outcomes.

Methods: Secondary analysis of the Brazilian Multicentre Study on Preterm Birth of 33 740 births in 20 maternity hospitals between April 2011 and July 2012. In accordance with a prototype concept based on maternal, fetal, and placental conditions, an adapted *k*-means model and fuzzy algorithm were used to identify clusters using predefined conditions. The mains outcomes were phenotype clusters and maternal and neonatal outcomes.

Results: Among 4150 PTBs, three clusters of PTB phenotypes were identified: women who had PTB without any predefined conditions; women with mixed conditions; and women who had pre-eclampsia, eclampsia, HELLP syndrome and fetal growth restriction. The prevalence of different preterm subtypes differed significantly in the three clusters, varying from 80.95% of provider-initiated PTBs in cluster 3–6.62% in cluster 1 ($P<0.001$). Although some maternal characteristics differed among the clusters, maternal and neonatal outcomes did not.

Conclusions: The analysis identified three clusters with distinct phenotypes. Women from the different clusters had different subtypes of PTB and maternal and pregnancy characteristics.

KEYWORDS

Cluster; *k*-means; Maternal outcomes; Neonatal outcomes; Phenotypes; Preterm birth

1 | INTRODUCTION

The limitations of current predictive algorithms reflect both the multifactorial nature of spontaneous preterm birth (sPTB) and the need to apply new strategies that can identify specific groups at risk. Known risk factors can play different roles in distinct subgroups of women.[1] There might be different pathways and complex interactions of conditions related to sPTB.

Specialists have proposed a new conceptual framework for PTB by selecting conditions that present in the index pregnancy, including maternal, fetal, and placental conditions that are not necessarily risk factors for PTB, but that are reasonably part of its pathways.[2,3] It is possible that there is not just one clinical phenotype related to PTB, and the identification of such phenotypes might shed light on the complex interactions among the underlying conditions related its occurrence.

A recent cluster analysis by a multi-ethnic international multicenter study showed that 30% of all cases of sPTB had no maternal, fetal, or placental conditions that might be related to its occurrence.[4] On the other hand, there were clusters characterized by conditions potentially sharing common severe maternal conditions with similar pathophysiological underlying conditions such as pre-eclampsia, third trimester bleeding and fetal growth restriction. In addition, it was possible to specify the most frequent clinical conditions related to its occurrence. Furthermore, not only were the predisposing causes shown to vary in the different clusters, but the maternal and neonatal outcomes were also distinct.[4]

A secondary analysis using a database of sPTB cases used a different clustering approach to establish nine clinical phenotypes with three levels of evidence for each phenotype.[5] After a hierarchical cluster analysis, PTB cases were grouped into five clusters characterized by different conditions such as maternal stress, premature rupture of membranes, familial factors, maternal morbidities, and multifactorial conditions. According to the study authors, women from the same cluster were more likely to share common causes and common genetic predispositions.[5]

Clustering analysis applied to PTB determinants is thus an innovative approach to identify groups of women who might require special attention, interventions, and surveillance depending on the conditions associated with the different subtypes of PTB and also the maternal and perinatal outcomes. This might be helpful for the identification of clinical phenotypes related to specific subtypes of PTB, and also facilitate studies of its determinants and associated outcomes, because the maternal clinical conditions can be identified by clinicians and healthcare providers during prenatal care.

The aim of the present study was therefore to perform a secondary analysis of The Brazilian Multicentre Study on Preterm Birth (EMIP) to identify whether there is a correlation between clustering of clinical, maternal, and fetal conditions and PTB subtypes, and to demonstrate maternal and neonatal outcomes related to the final clusters.

2 | MATERIALS AND METHODS

The present secondary cluster analysis was based on data from EMIP, a multicenter cross-sectional study with a nested case–control component of PTB conducted between April 1, 2011, and September 30, 2012, that collected comprehensive data related to the three subtypes

TABLE 1 Definition of maternal, fetal, and placental conditions potentially associated with preterm birth.

Condition	Definition
Maternal	
Extrauterine infection during pregnancy	Prenatal care chart or medical record of syphilis, tuberculosis, HIV, HPV, hepatitis, febrile diarrhea, pneumonia, sinusitis, toxoplasmosis, genital herpes, asymptomatic bacteriuria, cystitis, pyelonephritis or sepsis during pregnancy
Clinical chorioamnionitis	Medical record of clinical chorioamnionitis
Maternal chronic disease	Medical record of history of diabetes, HIV, chronic hypertension, hypo-/hyperthyroidism, nephropathy, sickle cell disease or other chronic anemia, cardiopathy, pneumopathy, epilepsy, lupus, other collagenosis, chronic gastrointestinal, psychiatric, neurologic or orthopedic disease, neoplasms, thrombosis, thrombophilia, or bariatric surgery
Pre-eclampsia/eclampsia/HELLP syndrome	Medical record of pre-eclampsia, eclampsia and/or HELLP syndrome
Fetal	
Antepartum stillbirth	Antepartum stillbirth (>22 gestational weeks) before or after hospital admission
FGR (suspicious or confirmed)	Suspicious or confirmed FGR, including cases with ultrasound scan estimated fetal weight <10th percentile during prenatal care and neonates with birthweight small for gestational age
Perinatal sepsis	Medical record of clinical or laboratory diagnosis of neonatal sepsis
Multiple pregnancy	Pregnancy with more than one live fetus after 12 gestational weeks
Fetal anomaly	Suspicious (ultrasound findings of fetal anomaly) or confirmed (after delivery) minor or major fetal anomaly
Placental	
Early bleeding	Reported bleeding before 13 gestational weeks
Mid/late pregnancy bleeding	Reported bleeding after 14 gestational weeks
None	PTB cases with none of the above conditions

Abbreviations: FGR, fetal growth restriction; HPV, human papilloma virus; PTB, preterm birth.

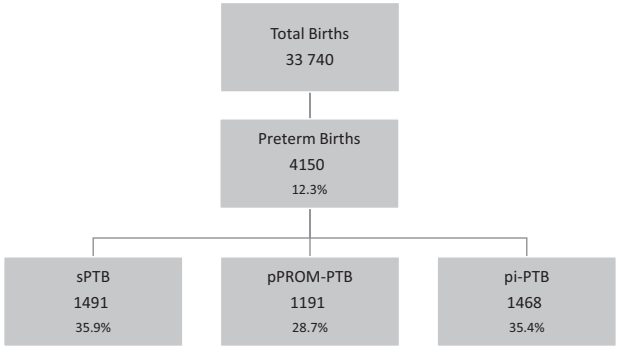

FIGURE 1 Flowchart showing the study population of the Brazilian Multicentre Study on Preterm Birth. The preterm birth subtypes were spontaneous preterm birth (sPTB); preterm birth due to preterm premature rupture of membranes (pPROM-PTB); and provider-initiated preterm birth (pi-PTB).

of PTB in 20 referral maternity hospitals in three regions of Brazil.[6–8] All participating women signed an informed consent form. The ethical principles stated in the Brazilian National Heath Council (Resolution CNS 466/12) were respected. The study also complied with the Declaration of Helsinki amended in Hong Kong in 1989. The original study was approved by the local institutional review board of the coordinating center, by each local institutional review board of all participating centers, and by the National Ethics Committee for Research.

The EMIP has been previously described.[6–8] In brief, it was a comprehensive observational study that identified all PTBs occurring in 20 referral facilities with more than 33 000 deliveries, and collected more than 300 variables related to potentially associated factors and maternal and neonatal outcomes. Information about medical history, sociodemographic status, and pregnancy, delivery, and postpartum details were retrospectively collected after delivery through an interview with the participating women and a review of hospital medical records including prenatal charts. Maternal and neonatal data were collected until either discharge or 40 days after delivery.

The present analysis used the concept framework and maternal, fetal, and placental conditions of Barros et al.,[3,4] which were defined as potential conditions that might be directly or indirectly related to the occurrence of PTB (Table 1). These conditions were used to establish the different preterm phenotypes.

Preterm birth was classified as one of three subtypes: spontaneous preterm birth (sPTB) due to spontaneous onset of labor; premature rupture of membranes leading to preterm birth (PROM-PTB); or provider-initiated preterm birth (pi-PTB) due to maternal and/or fetal conditions motivating preterm delivery.

The distribution of maternal and neonatal outcomes, including mode of delivery, gestational age category (extreme, moderate, and late preterm), Apgar score <7 at 5 minutes, admission to neonatal intensive care unit (NICU), neonatal near miss (based on birthweight below 1700 g, Apgar score below 7 at 5 minutes of life, and gestational age <33 weeks), and neonatal death before discharge was determined in the clusters. The distribution of some maternal and pregnancy characteristics in the PTB clusters was also determined. Adequacy of weight gain was categorized as insufficient, adequate, and excessive in accordance with the US Institute of Medicine definition for weekly rate of weight gain.[9]

Statistical analysis was conducted by using SAS version 9.4 (SAS Institute, Cary, NC, USA). A cluster analysis was conducted to identify clusters dependent on the predefined maternal, fetal and placental conditions listed in Table 1. A *k*-modes model, which is a variation of the *k*-means model for categoric variables, was applied to identify clusters from the predefined conditions using a fuzzy algorithm. The number of final clusters was determined by automatized methods (no predefined number of clusters was set). χ^2 test was used to evaluate differences in maternal and neonatal outcomes among the clusters. A *P* value of less than 0.05 was taken to indicate significance.

3 | RESULTS

Among a total of 33 740 births during the study period, the EMIP study collected data on 4150 PTBs, which were categorized as sPTB (n=1491), pPROM-PTB (n=1191), and pi-PTB (n=1468) (Fig. 1).

TABLE 2 Distribution of clusters of PTB phenotypes according to maternal, fetal, and neonatal conditions.

Cluster	No. (%) of PTBs (n=4150)	Main condition		Other frequent conditions[a]	
		Type	%	Type	%
1	650 (15.7)	None	100		
2	2319 (55.9)	Extrauterine infection	42.5	Maternal chronic disease	34.9
				Mid/late pregnancy bleeding	20.2
				Multiple pregnancy	15.6
				Clinical chorioamnionitis	7.5
				Antepartum stillbirth	6.3
3	1181 (28.4)	Pre-eclampsia/eclampsia/HELLP syndrome	85.8	Fetal growth restriction	32.2

Abbreviations: HELLP syndrome, hemolysis, elevated liver enzymes, low platelet count; PTB, preterm birth.
[a]Conditions with prevalence >30% or almost exclusively within that cluster (>80%).

TABLE 3 Distribution of maternal, fetal and placental conditions according to clusters of preterm birth phenotype.

Condition	Cluster 1 (n=650)	Cluster 2 (n=2319)	Cluster 3 (n=1181)
Extrauterine infection, no.	0	986	343
Row, %	0	74.19	25.81
Column, %	0	42.52	29.04
Clinical chorioamnionitis, no.	0	173	0
Row %	0	100.0	0
Column %	0	7.46	0
Maternal chronic disease, no.	0	809	222
Row, %	0	78.47	21.53
Column, %	0	34.89	18.80
Pre-eclampsia/eclampsia/ HELLP syndrome, no.	0	51	1013
Row, %	0	4.79	95.21
Column, %	0	2.20	85.77
Antepartum stillbirth, no.	0	147	17
Row, %	0	89.63	10.37
Column, %	0	6.34	1.44
Fetal growth restriction, no.	0	49	380
Row, %	0	11.42	88.58
Column, %	0	2.11	32.18
Perinatal sepsis, no.	0	564	212
Row, %	0	72.68	27.32
Column, %	0	24.32	17.95
Multiple pregnancy, no.	0	362	75
Row, %	0	82.84	17.16
Column, %	0	15.61	6.35
Fetal anomaly, no.	0	383	112
Row, %	0	77.37	22.63
Column, %	0	16.52	9.48
Early bleeding, no.	0	431	134
Row, %	0	76.28	23.72
Column, %	0	18.59	11.35
Mid/late pregnancy bleeding, no.	0	468	86
Row, %	0	84.48	15.52
Column, %	0	20.18	7.28
None, no.	650	0	0
Row, %	100	0	0
Column, %	100	0	0

Row, %: Distribution of women with a given condition in a particular cluster. Column, %: Distribution of different conditions in a particular cluster.

The 4150 cases of PTB were clustered into three groups according to the 12 predefined maternal, fetal, and placental conditions (Table 1). Not having any predefined condition was also considered to be a 'predefined condition'. The prevalence of the main condition and the next most frequent conditions in the three clusters are presented in Table 2.

Cluster 1 (n=650, 15.7%) was characterized by women who did not have any defined maternal, fetal, or placental conditions. Cluster 2 (n=2319, 55.9%) was characterized by the following set of conditions: 42.5% had extrauterine infection, 34.9% had maternal chronic disease, and approximately 20% had mid-late pregnancy bleeding. All women who had clinical chorioamnionitis, almost 90% who had antepartum stillbirth, and more than 80% who had multiple pregnancy were in cluster 2 (Tables 2 and 3). In cluster 3 (n=1181, 28.4%), 85.8% of women had pre-eclampsia, eclampsia, or HELLP syndrome (hemolysis, elevated liver enzymes, and low platelet count), and 32.2% had fetal growth restriction.

Table 3 summarizes the distribution of the 11 predefined conditions in the three clusters, including the prevalence and concentration of each given condition in the clusters. Although only 7.46% of women in cluster 2 had clinical chorioamnionitis, all women with clinical chorioamnionitis were contained in cluster 2.

The PTB subtype differed by cluster ($P<0.001$) (Table 4). More than 90% of women in cluster 1 had sPTB or pPROM-PTB. The proportion of women with pi-PTB was slightly higher in cluster 2 (20.22%) and much higher in cluster 3 (80.95%) as compared with cluster 1 (6.62%).

The maternal and neonatal outcomes did not differ among the three clusters (Table 5). Cesarean was the most prevalent mode of delivery, ranging from 52.7% to 55.0% of PTBs in the clusters.

The distribution of maternal and pregnancy characteristics in the three clusters was determined (Table 6). White race, obesity (body mass index, calculated as weight in kilograms divided by the square of height in meters, >25), excessive weight gain during pregnancy, and previous cesarean delivery were more prevalent in cluster 3 than in cluster 2, and more prevalent in cluster 2 than in cluster 1. None of the other characteristics examined differed among the clusters.

4 | DISCUSSION

The present analysis found that the 4150 PTBs of the EMIP study were clustered into three groups, which presented with very different clinical conditions (phenotypes). The first cluster had no associated

TABLE 4 Preterm birth subtypes according to preterm birth phenotype cluster.

PTB subtype	Cluster 1, n (%)	Cluster 2, n (%)	Cluster 3, n (%)	*P* value
sPTB	343 (52.77)	1018 (43.90)	130 (11.01)	<0.001
pPROM-PTB	264 (40.62)	832 (35.88)	95 (8.04)	
pi-PTB	43 (6.62)	469 (20.22)	956 (80.95)	
All cases	650 (100)	2319 (100)	100	

Abbreviations: sPTB, spontaneous preterm birth; pPROM-PTB, preterm birth due to preterm premature rupture of membranes; pi-PTB, provider-initiated preterm birth.

TABLE 5 Maternal and neonatal outcomes according to preterm birth phenotype clusters.

Maternal and neonatal outcomes	Cluster 1, n (%)	Cluster 2, n (%)	Cluster 3, n (%)	*P* value
Cesarean delivery	359 (55.0)	1223 (52.7)	637 (53.9)	0.528
GA at delivery				
<28 wk	50 (7.6)	171 (7.3)	87 (7.3)	0.959
<32 wk	132 (20.3)	494 (21.3)	254 (21.5)	0.822
34–36 wk	406 (62.4)	1452 (62.6)	743 (62.9)	0.977
Apgar score <7 at 5 min[a]	64 (10.0)	228 (10.0)	138 (11.8)	0.227
NICU stay >7 d[a]	195 (34.1)	681 (34.2)	318 (32.0)	0.492
Neonatal near miss[a]	230 (35.5)	771 (33.3)	389 (33.0)	0.510
Neonatal death before discharge[a]	60 (9.5)	190 (8.5)	82 (7.3)	0.257

Abbreviation: NICU, neonatal intensive care unit.
[a]Data were missing for 69 (Apgar score), 597 (NICU stay), 13 (neonatal near miss), and 198 (neonatal death before discharge) cases.

conditions; the second cluster had mixed conditions; and the third cluster was related to pre-eclampsia and fetal growth restriction. No differences in maternal or perinatal outcomes were observed among the clusters; regarding PTB subtype, however, the prevalence of pi-PTB was significantly higher in cluster 3 ($P<0.001$).

The study used an unsupervised data-driven cluster analysis, which meant that pre-clusters were not predefined and the initial number of clusters was not established. This approach enables a more genuine clustering of cases according to the predefined clinical conditions. The reproducibility of cluster analysis might depend on the dataset, and also on the availability of the defined clinical conditions. Nevertheless, it was considered that the selected clinical conditions are reproducible and commonly addressed in PTB studies, and are potentially available regardless of the setting or population.

The EMIP study followed standardized data collection protocols and several procedures to assure data quality.[10] Nevertheless, the present analysis has some limitations. First, there were no data on cervical length, a maternal condition that is highly associated with the occurrence of sPTB.[11] Second, it was an observational study with retrospective data collection after delivery for variables related to pregnancy. Therefore, the classification of some conditions was based only on self-report by the participating women or on medical records/prenatal charts, limiting the standardization and audit. Last, the definition of maternal chronic disease was based on different diseases that have potentially distinct effects on maternal and fetal health during pregnancy.

In the present analysis, the conceptual framework used by Barros et al. [4] was adapted to determine predefined conditions potentially associated with PTB. A much smaller number of clusters was identified as compared with Barros et al.'s study, indicating that the final number of clusters might depend both on the criteria for predefined conditions and on the clustering method. In the present analysis, the number of clusters was set by the model, avoiding external adjustments, whereas Barros et al.[4] preferred to use a two-step cluster analysis, which enabled the development of pre-clusters with adjustment of the final clusters. These different methodologic approaches might account for the different findings.

The new conceptual framework requires validation and possibly the inclusion of new conditions in the model. For instance, information on cervical length was unavailable for the analysis of Barros et al.,[4] and less than 5% of women in the present analysis had a cervical length measurement recorded at 20–24 weeks (data not shown). Cervical length is an important condition that should be addressed, in addition to maternal anthropometric status at the beginning of pregnancy.

The proportion of women in cluster 1 (with no predefined condition) is meaningful. Approximately 15% of all women with PTB did not have any of the 12 conditions potentially associated with PTB. The prevalence of women without any conditions was even higher (30%) in the multi-country population-based study of Barros et al.[4] Clinical and epidemiologic data seem to have limited performance in recognizing conditions related to the occurrence of PTB.[12,13]

The mechanisms of preterm and term labor are not completely understood. The proportion of women who had pi-PTB or sPTB/pPROM-PTB in the three clusters differed significantly; therefore, the clinical characteristics of women in the different clusters was analyzed in an effort to explain why. Women in cluster 1 had a low rate of pi-PTB. The absence of maternal morbidity or any other related conditions in this cluster might indicate that sPTB due to spontaneous onset of labor or pPROM often are present without a background of maternal obstetric conditions, confirming the great challenge in recognizing the mechanisms that lead to preterm labor or pPROM. Esplin et al.[5] performed a cluster analysis of 1028 women with PTB, reporting one cluster with a strong familial history of PTB that might have a genetic contribution based on insulin gene analysis. The identification of specific groups of women sharing common genetic and clinical conditions might provide better understanding of the complex interactions of different biologic systems (i.e., maternal, fetal, and placental) related to PTB.[2,3,14]

Cluster 2 was characterized by women with mixed conditions (extrauterine infection, maternal morbidity, clinical chorioamnionitis, vaginal bleeding during pregnancy, and multiple pregnancy). Although all of these conditions are known risk factors for PTB, it is difficult to determine the role of each in the occurrence of PTB in this cluster. Almost

TABLE 6 Maternal and pregnancy characteristics according to preterm birth phenotype clusters.

Characteristic	Cluster 1 (n=650)[a]	Cluster 2 (n=2319)[a]	Cluster 3 (n=1181)[a]	*P* value
Age, y[b]				0.8503
<19	126 (19.41)	496 (21.39)	242 (20.49)	
19–35	428 (65.95)	1488 (64.17)	764 (64.69)	
>35	95 (14.64)	335 (14.45)	175 (14.82)	
Ethnicity				0.0424
White	262 (40.31)	1023 (44.11)	548 (46.40)	
Non-white	388 (59.69)	1296 (55.89)	633 (53.60)	
Schooling, y[b]				0.1261
<12	125 (80.53)	1792 (78.67)	948 (81.51)	
≥12	125 (19.47)	486 (21.33)	215 (18.49)	
Family income, US $[b]				0.1829
<400	354 (60.62)	1307 (61.65)	636 (58.30)	
≥400	230 (39.38)	813 (38.35)	455 (41.70)	
Initial BMI[b]				<0.0001
<18.5	58 (10.74)	193 (9.53)	51 (4.99)	
18.5–25	343 (63.52)	1161 (57.31)	486 (47.51)	
>25	139 (25.74)	672 (33.17)	486 (47.51)	
Pregnancy weight gain[b]				<0.0001
Insufficient	191 (38.20)	575 (30.49)	172 (18.01)	
Adequate	83 (16.60)	246 (13.04)	112 (11.73)	
Excessive	226 (45.20)	1065 (56.47)	671 (70.26)	
Parity				
1	306 (47.08)	1130 (48.73)	539 (45.64)	0.2150
≥2	66 (10.15)	253 (10.91)	147 (12.45)	0.2530
Previous PTB[b]	124 (19.14)	433 (18.72)	255 (21.67)	0.1105
Previous SGA[b]	107 (16.72)	365 (15.90)	220 (18.66)	0.1191
Previous cesarean[b]	131 (20.15)	475 (20.49)	296 (25.06)	0.0046
Smoking, alcohol, or other drugs[b]	97 (14.99)	386 (16.79)	185 (15.74)	0.4812
Stress[b]	276 (42.86)	937 (40.85)	507 (43.19)	0.3536

Abbreviations: BMI, body mass index (calculated as weight in kilograms divided by the square of height in meters); PTB, preterm birth; SGA, small for gestational age.

[a]Values are given as mean ± SD or number (percentage).

[b]Data were missing information for 1 (maternal age), 67 (schooling), 355 (family income), 561 (BMI) 809 (pregnancy weight gain), 12 (previous PTB), 35 (SGA), 1 (previous cesarean), 29 (smoking etc.) and 38 (stress) women.

80% of women in the mixed-conditions cluster (cluster 2) had sPTB or pPROM-PTB, confirming that women with this subtype of PTB may have a multiplicity of conditions, which the present cluster analysis resolved into an inseparable group, in contrast to the findings of Barros et al.[4]

Not surprisingly, the cluster of women with pre-eclampsia, eclampsia, or HELLP syndrome also included fetal growth restriction as the second most frequent condition (cluster 3). Both conditions are "great obstetric syndromes" that are directly linked to ischemic placental disease and share common altered placentation mechanisms.[15,16] Hypertensive disorders and fetal growth restriction are the most important indications for pi-PTB due to maternal or fetal conditions,[8,17] which explains the high rates of pi-PTB in cluster 3. The prevalence of obesity and excessive weight gain during pregnancy was higher in cluster 3 than in the other clusters. Both conditions are considered risk factors for hypertensive disorders, but not for fetal growth restriction.[18] It is estimated that pre-eclampsia and fetal growth restriction account for only approximately 12% of ischemic placental disease in PTBs.[19] Although there is a concurrence of pre-eclampsia and fetal growth restriction, which are followed by poorer outcomes, the risk factors and conditions associated with each condition do not invariably overlap.[19,20]

Although the present analysis identified three clusters with very distinct clinical phenotypes, we consider that a clearer definition of the predefined conditions would provide better cluster resolution, considering that women were grouped into a very few number of clusters and one of them included multiple mixed conditions. For example, infectious diseases are underlying causes of PTB; however, the lack of details regarding the severity, treatment received, and moment when women were affected by infectious disease might have underestimated the association of such conditions with PTB. This marked condition was grouped in the same cluster as many other conditions (cluster 2). Rather than simply noting that women had an infectious

disease, taking into account the etiology, site, severity, and treatment received (or not) might be a better approach to explore this condition. Esplin et al.[5] proposed a score discriminating clinical conditions according the level of evidence associating them with PTB (i.e., possible, moderate, and strong). The idea is to refine the presentation of the clinical phenotype of each cluster. Data mining might be another helpful clustering technique to determine clinical conditions and the correspondent clusters. This approach is used to interpret "big" data in complex syndromes with multiple interaction systems as genome data.[21–24] The combination of phenotype clusters with biologic markers might be an innovative initiative to study and predict PTB.

In conclusion, three PTB clusters were identified with different phenotypes of women: those without any predefined conditions; those with mixed conditions; and those with hypertensive disorders in pregnancy and fetal growth restriction. Although the maternal and neonatal outcomes did not differ, women in the three clusters had different subtypes of preterm delivery. Standardized methods and larger datasets might provide more reliable and helpful findings to contribute to the study of PTB phenotypes.

BRAZILIAN MULTICENTRE STUDY ON PRETERM BIRTH STUDY GROUP

The other members of the Brazilian Multicentre Study on Preterm Birth Study Group were Ricardo P. Tedesco, Giuliane J. Lajos, Marcelo L. Nomura, Patricia M. Rehder, Tabata Z. Dias, Maria L. Costa, Samira M. Haddad, Sergio T. Marba, Ruth Guinsburg, Francisco E. Martinez, Vilma Zotarelli, Lucio T. Gurgel, Francisco E. Feitosa, George N. Chaves, Ana M. Porto, Isabela C. Coutinho, Antonio C. Barbosa Lima, Elias F. Melo Jr, Débora F. Leite, Melania M. Amorim, Adriana S.O. Melo, Fabiana O. Melo, Marília G. Martins, Marynea V. Nunes, Cláudio S. Paiva, Moises D. Lima, Djacyr M. Freire, Edson G. Tristão, Denis J. Nascimento, Carlos A. Menezes, Marcelo Aquino, Janete Vettorazzi, Cintia E. Senger, Augusta M.B. Assumpção, Marcela A. F. Guedes, Maria E. L. Moreira, Vera T. Borges, Nelson L. Maia Filho, Jacinta P. Mathias, Eduardo Souza, Ana C.P. Zamarian, Silvana M. Quintana, Patrícia P.S. Melli, Fátima A. Lotufo, Kaliane Uzilin, Elvira A. Zanette, Carla B. Andreucci, Tenilson A. Oliveira, Laércio R. Oliveira, Marcos A. N. Santos, Nelson Sass, Mirian R. F. Silveira, Pedro R. Coutinho, and Luciana Siqueira.

AUTHOR CONTRIBUTIONS

RTS contributed to data collection, study conception and design, and wrote the manuscript. JGC, RCP and RP contributed to project development, data collection, and study conception and design. PFO and CMS contributed to data analysis. The Brazilian Multicentre Study on Preterm Birth study group contributed to project development and data collection. All authors reviewed and approved the final manuscript.

ACKNOWLEDGMENTS

The study was supported by a Grant from the CNPq (Brazilian National Research Council), and by FAPESP (Foundation for Support to Research of the State of Sao Paulo). The funders have no role in the study design, data collection or analysis, decision to publish, or preparation of the manuscript. The current secondary analyses of EMIP data was designed, sponsored, and developed as part of the Preterm SAMBA study group (RTS, JGC, and RP), a collaborative study group on Preterm Birth Research jointly issued by the Brazilian National Research Council (CNPq) and the Bill and Melinda Gates Foundation (award 401636/2013-5, Grand Challenges Brazil: Reducing the burden of preterm birth, number 05/2013). The current findings are thus part of the scope of analyses in preterm birth of the Preterm SAMBA study group. The EMIP study database is not available in data repositories, but data can be provided under approval of the EMIP study group.

CONFLICTS OF INTEREST

The authors have no conflicts of interest.

REFERENCES

1. Morisaki N, Ganchimeg T, Ota E, et al. Maternal and institutional characteristics associated with the administration of prophylactic antibiotics for caesarean section: A secondary analysis of the World Health Organization Multicountry Survey on Maternal and Newborn Health. *BJOG*. 2014;121:66–75.
2. Goldenberg RL, Gravett MG, Iams J, et al. The preterm birth syndrome: Issues to consider in creating a classification system. *Am J Obstet Gynecol*. 2012;206:113–118.
3. Villar J, Papageorghiou AT, Knight HE, et al. The preterm birth syndrome: A prototype phenotypic classification. *Am J Obstet Gynecol*. 2012;206:119–123.
4. Barros FC, Papageorghiou AT, Victora CG, et al. The distribution of clinical phenotypes of preterm birth syndrome: Implications for prevention. *JAMA Pediatr*. 2015;169:220–229.
5. Esplin MS, Manuck TA, Varner MW, et al. Cluster analysis of spontaneous preterm birth phenotypes identifies potential associations among preterm birth mechanisms. *Am J Obstet Gynecol*. 2015;213:429.e1–429.e9.
6. Passini R, Tedesco RP, Marba ST, et al. Brazilian multicenter study on prevalence of preterm birth and associated factors. *BMC Pregnancy Childbirth*. 2010;10:22.
7. Passini R, Cecatti JG, Lajos GJ, et al. Brazilian multicentre study on preterm birth (EMIP): Prevalence and factors associated with spontaneous preterm birth. *PLoS ONE*. 2014;9:e109069.
8. Souza RT, Cecatti JG, Passini R, et al. The burden of provider-initiated preterm birth and associated factors: Evidence from the Brazilian Multicenter Study on Preterm Birth (EMIP). *PLoS ONE*. 2016;11:e0148244.
9. Institute of Medicine (IOM), National Research Council (NRC). *Weight Gain During Pregnancy: Reexamining the Guidelines*. Washington DC: The National Academies Press; 2009.
10. Lajos GJ, Tedesco RP, Passini R, et al. Methodological issues on planning and running the brazilian multicenter study on preterm birth. *ScientificWorldJournal*. 2015;2015:719104.
11. Berghella V, Palacio M, Ness A, et al. Cervical length screening for prevention of preterm birth in singleton pregnancy with threatened preterm labor: Systematic review and meta-analysis of randomized controlled trials using individual patient-level data. *Ultrasound Obstet Gynecol*. 2017;49:322–329.
12. Koullali B, Oudijk MA, Nijman TAJ, et al. Risk assessment and management to prevent preterm birth. *Semin Fetal Neonatal Med*. 2016;21:80–88.

13. Honest H, Hyde CJ, Khan KS. Prediction of spontaneous preterm birth: No good test for predicting a spontaneous preterm birth. *Curr Opin Obstet Gynecol*. 2012;24:422–433.
14. Goldenberg RL, Goepfert AR, Ramsey PS. Biochemical markers for the prediction of preterm birth. *Am J Obstet Gynecol*. 2005;192:S36–S46.
15. Di Renzo GC. The great obstetrical syndromes. *J Matern neonatal Med*. 2009;22:633–635.
16. Brosens I, Pijnenborg R, Vercruysse L, Romero R. The, "Great Obstetrical Syndromes" are associated with disorders of deep placentation. *Am J Obstet Gynecol*. 2011;204:193–201.
17. Morisaki N, Togoobaatar G, Vogel J, et al. Risk factors for spontaneous and provider-initiated preterm delivery in high and low Human Development Index countries: A secondary analysis of the World Health Organization Multicountry Survey on Maternal and Newborn Health. *BJOG*. 2014;121(Suppl):101–109.
18. Bartsch E, Medcalf KE, Park AL, et al. Clinical risk factors for pre-eclampsia determined in early pregnancy: Systematic review and meta-analysis of large cohort studies. *BMJ*. 2016;353:i1753.
19. Parker SE, Werler MM. Epidemiology of ischemic placental disease: A focus on preterm gestations. *Semin Perinatol*. 2014;38:133–138.
20. Ananth CV, Vintzileos AM. Ischemic placental disease: Epidemiology and risk factors. *Eur J Obstet Gynecol Reprod Biol*. 2011;159:77–82.
21. Beckmann JS, Lew D. Reconciling evidence-based medicine and precision medicine in the era of big data: Challenges and opportunities. *Genome Med*. 2016;8:134.
22. Oyelade J, Isewon I, Oladipupo F, et al. Clustering algorithms: Their application to gene expression data. *Bioinform Biol Insights*. 2016;10:237–253.
23. Wang L, Wang Y, Chang Q. Feature selection methods for big data bioinformatics: A survey from the search perspective. *Methods*. 2016;111:21–31.
24. Peek N, Holmes JH, Sun J. Technical challenges for big data in biomedicine and health: Data sources, infrastructure, and analytics. *Yearb Med Inform*. 2014;9:42–47.