The use of metabolomics for predicting spontaneous preterm birth in asymptomatic pregnant women: protocol for a systematic review and meta-analysis

SCHOLARONE™
Manuscripts

1

2

3   **The use of metabolomics for predicting spontaneous preterm birth in asymptomatic**

4

5   **pregnant women: protocol for a systematic review and meta-analysis**

6

7

8

9

10   Renato T. Souza (renatotsouzasp@gmail.com)

11

12   Rafael Bessa Freitas Galvão (rafaelbfg@gmail.com)

13

14   Débora Farias Batista Leite (deborafariasleite@gmail.com)

15

16   Renato Passini Jr (passini@caism.unicamp.br)

17

18

19   Philip N. Baker (philip.baker@leicester.ac.uk)

20

21   José Guilherme Cecatti (cecatti@unicamp.br)

22

23

24

25

26   Department of Obstetrics and Gynaecology, University of Campinas, Campinas, São Paulo,

27

28   Brazil.

29

30   Clinics Hospital of Federal University of Pernambuco, Recife, Pernambuco, Brazil.

31

32

33   College of Life Sciences, University of Leicester, England, United Kingdom.

34

35

36

37

41

42

43

44   Corresponding author:

45

46   Renato T Souza

47

48   101 Alexander Fleming, Cidade Universitária, campinas, São Paulo, Brazil.

49

50   ZIPCODE 13083-881

51

52

53   renatotsouzasp@gmail.com

54

55

56   Word count: 2,921.

57

58   1

59

25 **ABSTRACT**

26 **Introduction**: Preterm birth (PTB) is the leading cause of neonatal mortality and short- and

27 long-term morbidity. The aetiology and pathophysiology of spontaneous PTB are still

28 unclear, which makes the identification of reliable and accurate predictor markers more

29 difficult, particularly for unscreened or asymptomatic women. Metabolomics biomarkers

30 have been demonstrated to be potentially accurate biomarkers for many disorders with

31 complex mechanisms such as PTB. Therefore, we aim to perform a systematic review of

32 metabolomics markers associated with spontaneous PTB. Our research question is "What is

33 the performance of metabolomics for predicting spontaneous preterm birth in

34 asymptomatic pregnant women?"

35 **Methods and analysis**: We will focus on studies assessing metabolomics techniques for

36 predicting spontaneous preterm birth in asymptomatic pregnant women. We will conduct a

37 comprehensive systematic review of the literature from the last 10 years. Only

38 observational cohort and case-control studies will be included. Our search strategy will be

39 carried out by two independent reviewers, who will scan title and abstract before carrying

40 out a full review of the article. The scientific databases to be explored include PubMed,

41 MedLine, ScieLo, EMBASE, LILACS, Web of Science, Scopus, and others.

42 **Ethics and dissemination:** This systematic review protocol does not require ethical approval.

43 We intend to disseminate our findings in scientific peer-reviewed journal, the Preterm

44 SAMBA study open access website, specialists' conferences, and to our funding agencies.

45 Registration details: This protocol is registered in PROSPERO platform (code

46 CRD42018100172).

47 **Keywords**: preterm birth, spontaneous preterm birth, metabolomics, biomarkers,

48 prediction.

2

49 **Strengths and limitations of this study**

50 ● This systematic review protocol takes into account some important aspects regarding

51 conducting a systematic review about spontaneous preterm birth and metabolomics such as

52 the criteria used for defining spontaneous preterm birth, different population risk

53 stratification, method used to estimate gestational age, and metabolomics techniques

54 details.

55 ● Two independent reviewers are responsible for searching and selecting studies, as also

56 extracting data, and a third reviewer will resolve any disagreement.

57 ● If possible, proper statistical methods will be applied to investigate metabolomics

58 accuracy in predicting spontaneous preterm birth.

59

3

60   **INTRODUCTION**

61   Spontaneous preterm birth (sPTB) is the leading cause of perinatal mortality and short- and

62   long-term morbidity [1,2]. It is defined as birth that occurs before 37 weeks gestation due to

63   spontaneous onset of labour or preterm premature rupture of membranes (pPROM) [3,4].

64   Several pathways and mechanisms linked with preterm birth have been proposed including,

65   neuroendocrine, vascular, immune-inflammatory, and behavioural processes [5]. More

66   specifically, several markers associated with uterine distension/contraction, decidual

67   inflammation/infection and activation of hypothalamic–pituitary–adrenal axis had been

68   studied in the past decades [5,6]. However, no single marker or combination of markers has

69   been found to be accurate enough for predicting sPTB [7–10].

70   Preterm birth is a complex and multifactorial syndrome that possibly has a long pre-clinical

71   phase, maternal and fetal interactions, genetic and environmental influences, and adaptive

72   mechanisms [11,12]. These challenging aspects, and the presence of still unknown

73   underlying mechanisms, are the main limitations for the identification of an accurate

74   predictor for sPTB [13–15]. None of the predictors used in clinical practice, such as previous

75   history of preterm birth, infection (vaginal and urinary contaminants), fibronectin and

76   transvaginal ultrasonography cervical length demonstrated exceptional accuracy for

77   predicting spontaneous preterm birth [7]. An exploration of innovative approaches is

78   urgently required.

79   Metabolomics is the study of metabolites, through identification and quantification of low-

80   weight molecular particles, i.e. tens to hundreds thousands of intermediate products and

81   substrates of systems biology chemical reactions [16,17]. This novel approach has been

82   applied for identifying biomarkers and underlying biochemical pathways associated with

83   complex obstetrical syndromes as preeclampsia, fetal growth restriction, gestational

4

84   diabetes and preterm birth. In contrast to other "*Omics Sciences*" techniques, metabolomics

85   is more closely associated with the phenotype of the disease and might thus identify a more

86   robust and reliable set of predictors [18]. Importantly, implementing an adequate *Omics*

87   experimental design is crucial for metabolomics studies.

88   Using different baseline population (asymptomatic vs symptomatic or low- vs high-risk

89   women for developing sPTB), study designs (prospective cohorts, case-control or cross

90   sectional studies), sources of samples (amniotic fluid, vaginal fluid, blood, urine, hair, etc)

91   and the timing of sample collection each have significant effects on study findings and the

92   consequent interpretation and contribution to the current gap of knowledge [16].

93   Therefore, we aim to conduct a systematic review of the use of metabolomics biomarkers

94   for predicting spontaneous preterm birth in asymptomatic pregnant women. This protocol

95   describes the methods that will be applied in our systematic review.

96   **METHODS AND ANALYSIS**

97   The current systematic review proposal will be conducted, written and published following

98   the Preferred Reporting Items for Systematics Reviews and Meta-Analyses (PRISMA-P)

99   recommendations [19].   Also, it is properly registered at PROSPERO platform – code

100   CRD42018100172.

101   **Review question**

102   What is the performance of metabolomics for predicting spontaneous preterm birth in

103   asymptomatic pregnant women?

104   **Eligibility Criteria**

105   Original cohort or case-control studies involving asymptomatic pregnant women at the

106   moment of sample collection (exposure) and with samples analysed using metabolomics

5

107   techniques. Studies will be excluded if (1) they are cross-sectional studies, clinical trials,

108   editorials, letter to editors, case reports, expert opinions, commentaries, or any type of

109   review; (2) they describes only experimental studies with animals; or (3) they are duplicated

110   data (e.g. data published in conferences proceedings and, then, published again in scientific

111   journals). In this case, only the most complete publication will be considered, after

112   comparing and confirming that the same technique and metabolites were explored. Studies

113   published from 2008 to 2018 will be considered, and there will be no language restriction.

114   Before submitting this systematic review for publication, we will rerun the search strategy to

115   identify new studies that have been published after performing the first round of search.

116   **Participants**

117   The current review is interested in evaluating the performance of metabolomics biomarkers

118   for spontaneous preterm birth in asymptomatic women, which may contribute to clinical

119   practice, potentially providing information regarding onset of preterm labour. Nevertheless,

120   we aim to identify studies addressing only early predictors collected from asymptomatic

121   women (i.e. women who are in an early preclinical stage), which might contribute to a wider

122   window of opportunity for interventions and also to develop a widely reproducible

123   screening test. Asymptomatic pregnant women should not have regular uterine

124   tightening/contractions or signs of rupture of membranes (i.e. watery discharge). In

125   addition, the study should preferably have a standardized definition of spontaneous

126   preterm birth, the outcome of interest.

127   **Information Sources**

128   The search will be held in the following databases: PubMed, EMBASE, ProQuest, Scopus,

129   CINAHL, and Web of Science, BVS/BIREME, which includes the Latin American and

130   Caribbean Health Sciences Literature (LILACS), Medline and the Scientific Electronic Library

6

131 Online (Scielo). In addition, secondary sources of original studies will be explored such as

132 Google Scholar, hand-held searching of the reference list of eligible studies, conference

133 proceedings, and contact with authors when necessary.

134 **Search Strategy**

135 The following terms will be used in our search strategy for the different scientific databases:

136 (preterm birth, premature birth, premature infant, premature labor, extremely premature

137 infant, premature obstetric labor, spontaneous preterm birth, extreme preterm birth, late

138 preterm birth, moderate preterm birth, preterm premature rupture of membranes, preterm

139 delivery, PROM, sPTB, preterm PROM, pPROM, p-PROM) AND (metabolomic*,

140 metabonomic*, metabolit*, lipidomic*, H NMR, proton NMR, proton nuclear magnetic

141 resonance, liquid chromatogra*, gas chromatogra*, UPLC, ultra-performance liquid

142 chromatograph*, ultra performance liquid chromatograph*, HPLC, high performance liquid

143 chrromatograph*, high-performance liquid chrromatograph*) AND (pregnan*, antenat*,

144 ante nat*, prenat*, pre nat*). Respective adaptations in the syntax of search for each

145 database will be applied accordingly. No filters - such as "research in animal's models" and

146 "reviews" - will be used in our search strategy, as it will be excluded according to eligibility

147 criteria. The complete search strategy, including Boolean terms, is provided as

148 Supplementary Material.

149 **Data Management**

150 We will export search results to a reference manager (Mendeley®). Then, the following

151 information will be collected from each study using an appropriate form, which will be

152 entered in an Excel® spreadsheet: author's name, year of publication, country, study design,

153 number of participants with and without spontaneous preterm birth, type of metabolomics

154 analysis technique (liquid or gas chromatography, nuclear resonance), laboratory methods

7

155   for metabolites data acquiring (targeted or untargeted techniques, etc), subtype of preterm

156   birth (spontaneous or pPROM), number of fetuses (singleton vs multiple), gestational age

157   when samples were collected, source of samples (type/site of tissue), low or high-risk for

158   preterm birth (authors criteria used to define the population will be collected) and method

159   applied to estimate gestational age. If possible, additional variables related to spontaneous

160   preterm birth categories (delivery before 28 weeks and before 34 weeks) will be recorded

161   for secondary analyses. Finally, we will check the biochemical class of identified metabolites

162   in Human Metabolome Database (HMDB, version 4.0) to explore and synthetize whether

163   there are common biological pathways associated with spontaneous preterm birth [17].

164   **Selection Process**

165   Two independent reviewers (RTS and RBFG) will be responsible for screening and selecting

166   studies initially according to title or abstract. Full text of non-excluded studies will be read to

167   discriminate eligibility. A third reviewer (DFBL) will consider any disagreement; additional

168   reviewers (RPJ, PNB and JGC) will be responsible for supervising all steps and approving data

169   extraction.

170   **Data Collection Process**

171   We will extract search results to a reference manager where all studies will be stored. Then,

172   included studies will be placed in a new folder. Finally, we will manually extract data of

173   interest from these included studies to an Excel® file. Each reviewer will have their own

174   reference manager account, file and folder and discrepant results will be discussed together

175   with the third reviewer.

176   **Outcomes and Prioritization**

8

177 The primary outcome is spontaneous preterm birth, defined as any birth occurred before 37

178 weeks of gestation due to spontaneous onset of labor or preterm premature rupture of

179 membranes. Secondary outcomes are:

180   1) Spontaneous preterm birth before 28 weeks;

181   2) Spontaneous preterm birth before 34 weeks;

182 The capacity to predict different degrees of sPTB (categories of gestational age) is important

183 as the extreme (<28wks) and non-late preterm (<34wks) newborns have different adverse

184 outcomes compared to non-extreme (≥28wks) or late (≥34 wks) preterm newborns.

185 Ideally, the method of gestational age estimation should be clearly reported. For instance, it

186 can be reported as estimated by last menstrual period (LMP) and confirmed by an early

187 ultrasound or only by an early ultrasound when LMP is unknown/uncertain.

188 **Index test**

189 Metabolomics techniques to predict spontaneous preterm birth is the diagnostic test of

190 interest. Metabolomics is a technique to identify and quantify metabolites from biological

191 samples using different type of platforms/equipment. The most common platforms include

192 gas, liquid chromatography or ultra-performance liquid chromatography coupled to a mass

193 spectrometer or a proton nuclear magnetic resonance [20]. If possible, the performance of

194 each metabolomics techniques will be assessed through hierarchical summary receiver

195 operator characteristic curve (HSROC) (meta-analysis).

196 **Risk of Bias in individual Studies**

197 We will apply the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool [21]

198 to assess the risk of bias and applicability of primary diagnostic accuracy studies. Each study

199 will be classified as "low", "high" or "unclear" regarding risk of bias for each of the four

200 domains of QUADAS tool: Patient Selection, Index Test (metabolomics), Reference Standard

9

201 (occurrence of preterm birth), and Flow and Timing of participant's inclusion and follow-up.

202 For example, studies will be labelled as "low" risk of bias for Reference Standard when

203 definition of spontaneous preterm birth and gestational age estimation are clear; "high" risk

204 of bias would be considered when the moment of sample collection is not well described.

Data Synthesis

206 We will report details of identification, screening, eligibility and included studies using a flow

207 diagram, according to PRISMA recommendations [19]. Data from included studies will be

208 synthetized into tables according to the variables of interest. If possible, we will present

209 data meta-analysis according to study design, metabolomics technique and type of samples

210 analysed. We intend to perform subgroup analysis according to:

211 ● Different metabolomics methods applied: gas or liquid chromatography coupled

212 with mass spectrometry or proton nuclear magnetic resonance;

213 ● Singleton and multiple pregnancies;

214 ● Low-risk and high-risk women for developing preterm birth;

215 ● Subtype of preterm birth: Spontaneous preterm birth exclusively due to

216 spontaneous onset of labour with intact membranes or sPTB due to premature

217 rupture of membranes.

Potential anticipated limitations to this review

219 Firstly, although we have not considered any language restriction, we consider that there

220 might be a limitation in studies published entirely in non-English language. However, in the

221 last decade, more than 95% of scientific biomedical literature has been published in English

222 [22], then we consider this a minor selection bias. Secondly, we intend to stratify the groups

223 according to population risk. However, the characterization of low- or high-risk for

224 spontaneous preterm birth is controversial and lacks standardization, which might limit data

10

225 comparison and subgroup analysis. Finally, categorization of sPTB into spontaneous onset of

226 labour or pPROM is another topic of potential limitation - the recognition of the main initial

227 mechanism for preterm delivery might not always be possible. Even when specified, it might

228 provoke uncertainty and could limit further considerations regarding preterm phenotypes.

229 **Patient and Public Involvement**

230 Patients will not be directly involved in the study and no experience or direct impact from

231 their perspective can be discussed.

232 **ETHICS AND DISSEMINATION**

233 This systematic review does not require ethical approval from the Research Council or Ethics

234 board. We intend to disseminate our findings in scientific peer-reviewed journal, general

235 free access website of Preterm Screening and Metabolomics in Brazil and Auckland (Preterm

236 SAMBA) study, specialists' conferences, and to our funding agencies.

237 **DISCUSSION**

238 This systematic review will comprise current knowledge related with metabolomics in the

239 context of preterm birth prediction. Metabolomics science, a resourceful innovative field

240 that allows better understanding on pathophysiology of complex syndromes, may address

241 the main compounds associated with the spontaneous preterm delivery and, therefore,

242 motivate further researchers to validate early measurable predictors of preterm birth.

243 Metabolomics performance for predicting sPTB remains unclear and standardized and high-

244 quality studies are needed to clarify the clinical application of metabolites for predicting

245 sPTB. Nevertheless, metabolomics discovery studies commonly requires further validation

246 studies; reproducible methodology is crucial. This systematic review protocol will collate the

247 main potential early biomarkers, subgroup analysis and standardized definition for

11

248   spontaneous preterm birth to better understand metabolomics performance in predicting

249   sPTB and also to show its heterogeneity in terms of methodology (samples used,

250   metabolomics technique, definition of SPTB phenotype, etc). High performing predictors of

251   preterm birth will help combat this leading cause of neonatal mortality and morbidity.

252

253   **References**

254   1   Lawn JE, Cousens S, Zupan J. 4 million neonatal deaths: when? Where? Why? *Lancet*
255   2005;**365**:891–900. doi:10.1016/S0140-6736(05)71048-5

256   2   Howson CP, Kinney M V, McDougall L, *et al.* Born Too Soon: Preterm birth matters.
257   *Reprod Health* 2013;**10 Suppl 1**:S1. doi:10.1186/1742-4755-10-S1-S1

258   3   Blencowe H, Cousens S, Chou D, *et al.* Born Too Soon: The global epidemiology of 15
259   million preterm births. *Reprod Health* 2013;**10**:S2. doi:10.1186/1742-4755-10-S1-S2

260   4   Ananth C V, Vintzileos AM. Epidemiology of preterm birth and its clinical subtypes. *J
261   Matern Fetal Neonatal Med* 2006;**19**:773–82. doi:10.1080/14767050600965882

262   5   Behrman R, ButlerAS, editors. *Institute of Medicine (IOM). Preterm Birth: Causes,
263   Consequences, and Prevention.* Washington, D.C.: : National Academies Press 2007.
264   doi:10.17226/11622

265   6   Manuck TA, Esplin MS, Biggio J, *et al.* The phenotype of spontaneous preterm birth:
266   application of a clinical phenotyping tool. *Am J Obstet Gynecol* Published Online First:
267   February 2015. doi:10.1016/j.ajog.2015.02.010

268   7   Honest H, Hyde CJ, Khan KS. Prediction of spontaneous preterm birth: no good test
269   for predicting a spontaneous preterm birth. *Curr Opin Obstet Gynecol* 2012;**24**:422–
270   33. doi:10.1097/GCO.0b013e328359823a

271   8   Conde-Agudelo A, Papageorghiou A, Kennedy S, *et al.* Novel biomarkers for the
272   prediction of the spontaneous preterm birth phenotype: a systematic review and
273   meta-analysis. *BJOG An Int J Obstet Gynaecol* 2011;**118**:1042–54. doi:10.1111/j.1471-
274   0528.2011.02923.x

275   9   Hee L. Likelihood ratios for the prediction of preterm delivery with biomarkers. *Acta
276   Obstet Gynecol Scand* 2011;**90**:1189–99. doi:10.1111/j.1600-0412.2011.01187.x

277   10   Goldenberg RL, Iams JD, Mercer BM, *et al.* The Preterm Prediction Study: toward a
278   multiple-marker test for spontaneous preterm birth. *Am J Obstet Gynecol*
279   2001;**185**:643–51. doi:10.1067/mob.2001.116752

280   11   Di Renzo GC. The great obstetrical syndromes. *J Matern neonatal Med* 2009;**22**:633–

12

281   5. doi:10.1080/14767050902866804

282   12   Brosens I, Pijnenborg R, Vercruysse L, *et al.* The 'Great Obstetrical Syndromes' are
283   associated with disorders of deep placentation. *Am J Obstet Gynecol* 2011;**204**:193–
284   201. doi:10.1016/j.ajog.2010.08.009

285   13   Practice bulletin no. 130: prediction and prevention of preterm birth. *Obstet Gynecol*
286   2012;**120**:964–73. doi:10.1097/AOG.0b013e3182723b1b

287   14   Goldenberg RL, Culhane JF, Iams JD, *et al.* Epidemiology and causes of preterm birth.
288   *Lancet* 2008;**371**:75–84. doi:10.1016/S0140-6736(08)60074-4

289   15   Honest H, Hyde CJ, Khan KS. Prediction of spontaneous preterm birth. *Curr Opin*
290   *Obstet Gynecol* 2012;**24**:422–33. doi:10.1097/GCO.0b013e328359823a

291   16   Horgan RP, Clancy OH, Myers JE, *et al.* An overview of proteomic and metabolomic
292   technologies and their application to pregnancy research. *BJOG* 2009;**116**:173–81.
293   doi:10.1111/j.1471-0528.2008.01997.x

294   17   Wishart DS, Feunang YD, Marcu A, *et al.* HMDB 4.0: the human metabolome database
295   for 2018. *Nucleic Acids Res* 2018;**46**:D608–17. doi:10.1093/nar/gkx1089

296   18   Dettmer K, Hammock BD. Metabolomics--a new exciting field within the omics
297   sciences. *Environ Health Perspect* 2004;**112**:A396-7.

298   19   Shamseer L, Moher D, Clarke M, *et al.* Preferred reporting items for systematic review
299   and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ*
300   2015;**350**:g7647. doi:10.1136/BMJ.G7647

301   20   Zhang A, Sun H, Wang P, *et al.* Modern analytical techniques in metabolomics
302   analysis. *Analyst* 2012;**137**:293–300. doi:10.1039/c1an15605e

303   21   Whiting PF, Rutjes AWS, Westwood ME, *et al.* QUADAS-2: A Revised Tool for the
304   Quality Assessment of Diagnostic Accuracy Studies. *Ann Intern Med* 2011;**155**:529.
305   doi:10.7326/0003-4819-155-8-201110180-00009

306   22   Rosselli D. The language of biomedical sciences. *Lancet* 2016;**387**:1720–1.
307   doi:10.1016/S0140-6736(16)30259-8

308

309   **Author´s Contributions**

310   RTS and RBFG will conduct the systematic review as independent first reviewers. JGC, RPJ

311   and DFBL will decide about conflicting decisions regarding papers selections. PNB, RPJ and

13

312 JGC participated in the systematic review conception, methodology and framework,

313 together will all the others co-authors.

Competing interests

323 All authors are carrying original research about metabolomics and presenting conferences

324 about this topic, including spontaneous preterm birth, preeclampsia, gestational diabetes

325 mellitus and fetal growth restriction. Philip N Baker is principal investigator of Metabolomics

326 Diagnostics Ltd, a company dedicated to develop innovative screening tests using

327 metabolomics technology.

Ethics approval and consent to participate

14

333   This systematic review does not require ethical approval from the Research Council or Ethics

334   board.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Seacrh strategy: #1 AND #2 AND #3

1 (OR for each term)	preterm birth
	premature birth
	premature infant
	premature labor
	extremely premature infant
	premature obstetric labor
	spontaneous preterm birth
	extreme preterm birth
	late preterm birth
	moderate preterm birth
	preterm premature rupture of membranes
	preterm delivery
	PROM
	sPTB
	preterm PROM
	pPROM
	p-PROM

2 (OR for each term)	metabolomic*
	metabonomic*
	metabolit*
	lipidomic*
	H NMR
	proton NMR
	proton nuclear magnetic resonance
	liquid chromatogra*
	UPLC
	ultra-performance liquid chromatograph*
	ultra performance liquid chromatograph*
	HPLC
	high performance liquid chromatograph*
	high-performance liquid chromatograph*

3 (OR for each term)	pregnan*
	antenat*
	ante nat*
	prenat*
	pre nat*

1

PRISMA-P 2015 Checklist

This checklist has been adapted for use with systematic review protocol submissions to BioMed Central journals from Table 3 in Moher D et al**:** Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews* 2015 **4**:1

An Editorial from the Editors-in-Chief of *Systematic Reviews* details why this checklist was adapted - Moher D, Stewart L & Shekelle P**:** Implementing PRISMA-P: recommendations for prospective authors. *Systematic Reviews* 2016 **5**:15

Section/topic	#	Checklist item	Information reported		Line number(s)
			Yes	No	
ADMINISTRATIVE INFORMATION					
Title					
Identification	1a	Identify the report as a protocol of a systematic review	x	☐	2
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	☐	☐	n/a
Registration	2	If registered, provide the name of the registry (e.g., PROSPERO) and registration number in the Abstract	x	☐	48-49
Authors					
Contact	3a	Provide name, institutional affiliation, and e-mail address of all protocol authors; provide physical mailing address of corresponding author	x	☐	4-23
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review	x	☐	283-287
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments	☐	☐	n/a
Support					
Sources	5a	Indicate sources of financial or other support for the review	x	☐	275-282
Sponsor	5b	Provide name for the review funder and/or sponsor	x	☐	275-282
Role of sponsor/funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	x	☐	282
INTRODUCTION					
Rationale	6	Describe the rationale for the review in the context of what is already known	x	☐	51-87
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to	x	☐	94-95

2

Section/topic	#	Checklist item	Information reported		Line number(s)
			Yes	No	
		participants, interventions, comparators, and outcomes (PICO)			
METHODS					
Eligibility criteria	8	Specify the study characteristics (e.g., PICO, study design, setting, time frame) and report characteristics (e.g., years considered, language, publication status) to be used as criteria for eligibility for the review	x	☐	96-107
Information sources	9	Describe all intended information sources (e.g., electronic databases, contact with study authors, trial registers, or other grey literature sources) with planned dates of coverage	x	☐	119-125
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	x	☐	127-136
STUDY RECORDS					
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	x	☐	142-155
Selection process	11b	State the process that will be used for selecting studies (e.g., two independent reviewers) through each phase of the review (i.e., screening, eligibility, and inclusion in meta-analysis)	x	☐	157-161
Data collection process	11c	Describe planned method of extracting data from reports (e.g., piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	x	☐	163-167
Data items	12	List and define all variables for which data will be sought (e.g., PICO items, funding sources), any pre-planned data assumptions and simplifications	x	☐	142-155
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	x	☐	169-179
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	x	☐	189-196
DATA					
Synthesis	15a	Describe criteria under which study data will be quantitatively synthesized	x	☐	198-209
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data, and methods of combining data from studies, including any planned exploration of consistency (e.g., I^2, Kendall's tau)	x	☐	185-187
	15c	Describe any proposed additional analyses (e.g., sensitivity or subgroup analyses, meta-regression)	x	☐	185-187

3

Section/topic	#	Checklist item	Information reported		Line number(s)
			Yes	No	
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	x	☐	198-202
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (e.g., publication bias across studies, selective reporting within studies)	☐	☐	n/a
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (e.g., GRADE)	☐	☐	n/a

The use of metabolomics for predicting spontaneous preterm birth in asymptomatic pregnant women: protocol for a systematic review and meta-analysis

SCHOLARONE™
Manuscripts

1 **The use of metabolomics for predicting spontaneous preterm birth in asymptomatic**

2 **pregnant women: protocol for a systematic review and meta-analysis**

3

4 Renato T. Souza [1] (renatotsouzasp@gmail.com)

5 Rafael Bessa Freitas Galvão [1] (rafaelbfg@gmail.com)

6 Débora Farias Batista Leite [1,2] (deborafariasleite@gmail.com)

7 Renato Passini Jr [1] (passini@caism.unicamp.br)

8 Philip N. Baker [3] (philip.baker@leicester.ac.uk)

9 José Guilherme Cecatti [1] (cecatti@unicamp.br)

10

11 [1] Department of Obstetrics and Gynaecology, University of Campinas, Campinas, São Paulo,

12 Brazil.

13 [2] Clinics Hospital of Federal University of Pernambuco, Recife, Pernambuco, Brazil.

14 [3] College of Life Sciences, University of Leicester, England, United Kingdom.

15

16 This systematic review protocol is registered at PROSPERO platform (CRD42018100172). Available

17 from: http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42018100172.

18

19 Corresponding author:

20 José Guilherme Cecatti

21 101 Alexander Fleming, Cidade Universitária, campinas, São Paulo, Brazil.

22 ZIPCODE 13083-881

23 cecatti@unicamp.br

24 Word count: 2,233.

25 **ABSTRACT**

26 **Introduction**: Preterm birth (PTB) is the leading cause of neonatal mortality and short- and

27 long-term morbidity. The aetiology and pathophysiology of spontaneous PTB are still

28 unclear, which makes the identification of reliable and accurate predictor markers more

29 difficult, particularly for unscreened or asymptomatic women. Metabolomics biomarkers

30 have been demonstrated to be potentially accurate biomarkers for many disorders with

31 complex mechanisms such as PTB. Therefore, we aim to perform a systematic review of

32 metabolomics markers associated with spontaneous PTB. Our research question is "What is

33 the performance of metabolomics for predicting spontaneous preterm birth in

34 asymptomatic pregnant women?"

35 **Methods and analysis**: We will focus on studies assessing metabolomics techniques for

36 predicting spontaneous preterm birth in asymptomatic pregnant women. We will conduct a

37 comprehensive systematic review of the literature from the last 10 years. Only

38 observational cohort and case-control studies will be included. Our search strategy will be

39 carried out by two independent reviewers, who will scan title and abstract before carrying

40 out a full review of the article. The scientific databases to be explored include PubMed,

41 MedLine, ScieLo, EMBASE, LILACS, Web of Science, Scopus, and others.

42 **Ethics and dissemination:** This systematic review protocol does not require ethical approval.

43 We intend to disseminate our findings in scientific peer-reviewed journal, the Preterm

44 SAMBA study open access website, specialists' conferences, and to our funding agencies.

45 Registration details: This protocol is registered in PROSPERO platform (code

46 CRD42018100172).

47 **Keywords**: preterm birth, spontaneous preterm birth, metabolomics, biomarkers,

48 prediction.

2

Strengths and limitations of this study

- This systematic review protocol takes into account some important aspects regarding conducting a systematic review about spontaneous preterm birth and metabolomics such as the criteria used for defining spontaneous preterm birth, different population risk stratification, method used to estimate gestational age, and metabolomics techniques details.

- Two independent reviewers are responsible for searching and selecting studies, as also extracting data, and a third reviewer will resolve any disagreement.

- If possible, proper statistical methods will be applied to investigate metabolomics accuracy in predicting spontaneous preterm birth.

- Possible limitations to this review include the different criteria applied for defining spontaneous preterm birth, and the diverse population risk stratification.

3

62 **INTRODUCTION**

63 Spontaneous preterm birth (sPTB) is the leading cause of perinatal mortality and short- and

64 long-term morbidity [1,2]. It is defined as birth that occurs before 37 weeks gestation due to

65 spontaneous onset of labour or preterm premature rupture of membranes (pPROM) [3,4].

66 Several pathways and mechanisms linked with preterm birth have been proposed including,

67 neuroendocrine, vascular, immune-inflammatory, and behavioural processes [5]. More

68 specifically, several markers associated with uterine distension/contraction, decidual

69 inflammation/infection and activation of hypothalamic–pituitary–adrenal axis had been

70 studied in the past decades [5,6]. However, no single marker or combination of markers has

71 been found to be accurate enough for predicting sPTB [7–10]. History of previous preterm

72 birth, cervical length at second trimester and cervico-vaginal fetal fibronectin are the most

73 promising clinical tests for predicting spontaneous preterm, but they seem not to be

74 clinically useful for asymptomatic women. Sensitivity of short cervical length (<25mm) and

75 high cervico-vaginal fFN (>50ng/ml) are around 33-36% and 46%, respectively [11–13].

76 Preterm birth is a complex and multifactorial syndrome that possibly has a long pre-clinical

77 phase, maternal and fetal interactions, genetic and environmental influences, and adaptive

78 mechanisms [14,15]. These challenging aspects, and the presence of still unknown

79 underlying mechanisms, are the main limitations for the identification of an accurate

80 predictor for sPTB [16–18]. None of the predictors used in clinical practice, such as previous

81 history of preterm birth, infection (vaginal and urinary contaminants), fibronectin and

82 transvaginal ultrasonography cervical length demonstrated exceptional accuracy for

83 predicting spontaneous preterm birth [7]. An exploration of innovative approaches is

84 urgently required.

4

85 Metabolomics is the study of metabolites, through identification and quantification of low-

86 weight molecular particles, i.e. tens to hundreds thousands of intermediate products and

87 substrates of systems biology chemical reactions [19,20]. This novel approach has been

88 applied for identifying biomarkers and underlying biochemical pathways associated with

89 complex obstetrical syndromes as preeclampsia, fetal growth restriction, gestational

90 diabetes and preterm birth. In contrast to other "*Omics Sciences*" techniques, metabolomics

91 is more closely associated with the phenotype of the disease and might thus identify a more

92 robust and reliable set of predictors [21]. Importantly, implementing an adequate *Omics*

93 experimental design is crucial for metabolomics studies. Using different baseline population

94 (asymptomatic vs symptomatic or low- vs high-risk women for developing sPTB), study

95 designs (prospective cohorts, case-control or cross sectional studies), sources of samples

96 (amniotic fluid, vaginal fluid, blood, urine, hair, etc) and the timing of sample collection each

97 have significant effects on study findings and the consequent interpretation and

98 contribution to the current gap of knowledge [19].

99 Different reviews collating scientific knowledge regarding preterm birth

100 biomarkers/predictors has been conducted. Different methodology approaches has been

101 applied so far, including narrative, systematic and umbrella reviews, a more comprehensive

102 review that includes not only original studies but also other reviews [7,19,22,23]. At the best

103 of our knowledge, there is no systematic review on metabolomics markers. Therefore, we

104 aim to conduct a systematic review of original studies investigating the use of metabolomics

105 biomarkers for predicting spontaneous preterm birth in asymptomatic pregnant women.

106 This protocol describes the methods that will be applied in our systematic review.

107 **METHODS AND ANALYSIS**

108 The current systematic review proposal will be conducted, written and published following

109 the Preferred Reporting Items for Systematics Reviews and Meta-Analyses (PRISMA-P)

110 recommendations [24]. Also, it is properly registered at PROSPERO platform – code

111 CRD42018100172.

112 **Review question**

113 What is the performance of metabolomics for predicting spontaneous preterm birth in

114 asymptomatic pregnant women?

115 **Eligibility Criteria**

116 Original cohort or case-control studies involving asymptomatic pregnant women at the

117 moment of sample collection (exposure) and with samples analysed using metabolomics

118 techniques. Studies will be excluded if (1) they are cross-sectional studies, clinical trials,

119 editorials, letter to editors, case reports, expert opinions, commentaries, or any type of

120 review; (2) they describes only experimental studies with animals; or (3) they are duplicated

121 data (e.g. data published in conferences proceedings and, then, published again in scientific

122 journals). In this case, only the most complete publication will be considered, after

123 comparing and confirming that the same technique and metabolites were explored. Studies

124 published from 2008 to 2018 will be considered, and there will be no language restriction.

125 Before submitting this systematic review for publication, we will rerun the search strategy to

126 identify new studies that have been published after performing the first round of search.

127 **Participants**

128 The current review is interested in evaluating the performance of metabolomics biomarkers

129 for spontaneous preterm birth in asymptomatic pregnant women, which may contribute to

130 clinical practice, potentially providing information regarding onset of preterm labour.

131 Nevertheless, we aim to identify studies addressing only early predictors collected from

132   women who are in an early preclinical stage, which might contribute to a wider window of

133   opportunity for interventions and also to develop a widely reproducible screening test.

134   Asymptomatic pregnant women should not have regular uterine tightening/contractions or

135   signs of rupture of membranes (i.e. watery discharge). In addition, the study should

136   preferably have a standardized definition of spontaneous preterm birth, the outcome of

137   interest.

Information Sources

139   The search will be held in the following databases: PubMed, EMBASE, Scopus, CINAHL, and

140   Web of Science, BVS/BIREME, which includes the Latin American and Caribbean Health

141   Sciences Literature (LILACS), Medline and the Scientific Electronic Library Online (Scielo). In

142   addition, secondary sources of original studies will be explored such as Google Scholar,

143   hand-held searching of the reference list of eligible studies, conference proceedings, and

144   contact with authors when necessary.

Search Strategy

146   The following terms will be used in our search strategy for the different scientific databases:

147   (preterm birth, premature birth, premature infant, premature labor, extremely premature

148   infant, premature obstetric labor, spontaneous preterm birth, extreme preterm birth, late

149   preterm birth, moderate preterm birth, preterm premature rupture of membranes, preterm

150   delivery, PROM, sPTB, preterm PROM, pPROM, p-PROM) AND (metabolomic*,

151   metabonomic*, metabolit*, lipidomic*, H NMR, proton NMR, proton nuclear magnetic

152   resonance, liquid chromatogra*, gas chromatogra*, UPLC, ultra-performance liquid

153   chromatograph*, ultra-performance liquid chromatograph*, HPLC, high performance liquid

154   chrromatograph*, high-performance liquid chrromatograph*) AND (pregnan*, antenat*,

155   ante nat*, prenat*, pre nat*) (Supplementary Material). Respective adaptations in the

7

156 syntax of search for each database will be applied accordingly. No filters - such as "research

157 in animal's models" and "reviews" - will be used in our search strategy, as it will be excluded

158 according to eligibility criteria. The complete search strategy, including Boolean terms, is

159 provided as Supplementary Material.

160 **Data Management**

161 We will export search results to a reference manager (Mendeley®). Then, the following

162 information will be collected from each study using an appropriate form, which will be

163 entered in an Excel® spreadsheet: author's name, year of publication, country, study design,

164 number of participants with and without spontaneous preterm birth, type of metabolomics

165 analysis technique (liquid or gas chromatography, nuclear resonance), laboratory methods

166 for metabolites data acquiring (targeted or untargeted techniques, etc), subtype of preterm

167 birth (spontaneous preterm labour or pPROM), number of fetuses (singleton vs multiple),

168 gestational age when samples were collected, source of samples (type/site of tissue), low or

169 high-risk for preterm birth (authors criteria used to define the population will be collected)

170 and method applied to estimate gestational age. If possible, additional variables related to

171 spontaneous preterm birth categories (delivery before 28 weeks and before 34 weeks) will

172 be recorded for secondary analyses. Original authors will be contacted to clarify data, when

173 needed. Finally, we will check the biochemical class of identified metabolites in Human

174 Metabolome Database (HMDB, version 4.0) to explore and synthetize whether there are

175 common biological pathways associated with spontaneous preterm birth [20].

176 **Selection Process**

177 Two independent reviewers (RTS and RBFG) will be responsible for screening and selecting

178 studies initially according to title or abstract. Both researchers will read the full text of non-

179 excluded studies to discriminate eligibility. A third reviewer (DFBL) will consider any

8

180   disagreement; additional reviewers (RPJ, PNB and JGC) will be responsible for supervising all

181   steps and approving data extraction.

182   **Data Collection Process**

183   We will extract search results to a reference manager where all studies will be stored. Then,

184   included studies will be placed in a new folder. Finally, we will manually extract data of

185   interest from these included studies to an Excel® file. Each reviewer will have their own

186   reference manager account, file and folder and discrepant results will be discussed together

187   with the third reviewer.

188   **Outcomes and Prioritization**

189   The primary outcome is spontaneous preterm birth, defined as any birth occurred before 37

190   weeks of gestation due to spontaneous onset of labor or preterm premature rupture of

191   membranes. Secondary outcomes are:

192   1. Spontaneous preterm birth before 28 weeks;

193   2. Spontaneous preterm birth before 32 weeks;

194   3. Spontaneous preterm birth before 34 weeks;

195   The capacity to predict different degrees of sPTB (categories of gestational age) is important

196   as the extreme (<28wks), moderate (<32 weeks) and non-late preterm (<34wks) newborns

197   have different adverse outcomes compared to non-extreme (≥28wks); non-moderate

198   (≥32wks) or late (≥34 wks) preterm newborns.

199   Ideally, the method of gestational age estimation should be clearly reported. For instance, it

200   can be reported as estimated by last menstrual period (LMP) and confirmed by an early

201   ultrasound or only by an early ultrasound when LMP is unknown/uncertain.

202   **Index test**

203 Metabolomics techniques to predict spontaneous preterm birth is the diagnostic test of

204 interest. Metabolomics is a technique to identify and quantify metabolites from biological

205 samples using different type of platforms/equipment. The most common platforms include

206 gas, liquid chromatography or ultra-performance liquid chromatography coupled to a mass

207 spectrometer or a proton nuclear magnetic resonance [25]. If possible, the performance of

208 each metabolomics techniques will be assessed through hierarchical summary receiver

209 operator characteristic curve (HSROC) (meta-analysis).

210 **Risk of Bias in individual Studies**

211 We will apply the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool [26]

212 to assess the risk of bias and applicability of primary diagnostic accuracy studies. Each study

213 will be classified as "low", "high" or "unclear" regarding risk of bias for each of the four

214 domains of QUADAS tool: Patient Selection, Index Test (metabolomics), Reference Standard

215 (occurrence of preterm birth), and Flow and Timing of participant's inclusion and follow-up.

216 For example, studies will be labelled as "low" risk of bias for Reference Standard when

217 definition of spontaneous preterm birth and gestational age estimation are clear; "high" risk

218 of bias would be considered when the moment of sample collection is not well described.

219 **Data Synthesis**

220 We will report details of identification, screening, eligibility and included studies using a flow

221 diagram, according to PRISMA recommendations [24]. Data from included studies will be

222 synthetized into tables according to the variables of interest. If possible, we will present

223 data meta-analysis according to study design, metabolomics technique and type of samples

224 analysed. We intend to perform subgroup analysis according to:

225 • Different metabolomics methods applied: gas or liquid chromatography coupled

226 with mass spectrometry or proton nuclear magnetic resonance;

10

227   ● Singleton and multiple pregnancies;

228   ● Low-risk and high-risk women for developing preterm birth;

229   ● Subtype of preterm birth: Spontaneous preterm birth exclusively due to

230   spontaneous onset of labour with intact membranes or sPTB due to premature

231   rupture of membranes.

232 Heterogeneity will be assessed by Cochran's Q, Hotelling's T-squared (τ^2) and I^2 tests. Funnel

233 plots and sensitivity and cumulative analyses will be applied for detection of temporal

234 trends and publication bias.

235 Potential anticipated limitations to this review

236 Firstly, although we have not considered any language restriction, we consider that there

237 might be a limitation in studies published entirely in non-English language. However, in the

238 last decade, more than 95% of scientific biomedical literature has been published in English

239 [27], then we consider this a minor selection bias. Secondly, we intend to stratify the groups

240 according to population risk. However, the characterization of low- or high-risk for

241 spontaneous preterm birth is controversial and lacks standardization, which might limit data

242 comparison and subgroup analysis. Finally, categorization of sPTB into spontaneous onset of

243 labour or pPROM is another topic of potential limitation - the recognition of the main initial

244 mechanism for preterm delivery might not always be possible. Even when specified, it might

245 provoke uncertainty and could limit further considerations regarding preterm phenotypes.

246 In addition, another limitation is that individual patient data will not be collected.

247 Patient and Public Involvement

248 Patients will not be directly involved in the study and no experience or direct impact from

249 their perspective can be discussed.

250

11

251 **ETHICS AND DISSEMINATION**

252 This systematic review does not require ethical approval from the Research Council or Ethics

253 board. We intend to disseminate our findings in scientific peer-reviewed journal, general

254 free access website of Preterm Screening and Metabolomics in Brazil and Auckland (Preterm

255 SAMBA) study, specialists' conferences, and to our funding agencies.

256 **DISCUSSION**

257 This systematic review will comprise current knowledge related with metabolomics in the

258 context of preterm birth prediction. Metabolomics science, a resourceful innovative field

259 that allows better understanding on pathophysiology of complex syndromes, may address

260 the main compounds associated with the spontaneous preterm delivery and, therefore,

261 motivate further researchers to validate early measurable predictors of preterm birth.

262 Metabolomics performance for predicting sPTB remains unclear and standardized and high-

263 quality studies are needed to clarify the clinical application of metabolites for predicting

264 sPTB. Nevertheless, metabolomics discovery studies commonly requires further validation

265 studies; reproducible methodology is crucial. This systematic review protocol will collate the

266 main potential early biomarkers, subgroup analysis and standardized definition for

267 spontaneous preterm birth to better understand metabolomics performance in predicting

268 sPTB and also to show its heterogeneity in terms of methodology (samples used,

269 metabolomics technique, definition of SPTB phenotype, etc). High performing predictors of

270 preterm birth will help combat this leading cause of neonatal mortality and morbidity.

271

References

1 Lawn JE, Cousens S, Zupan J. 4 million neonatal deaths: when? Where? Why? *Lancet* 2005;**365**:891–900. doi:10.1016/S0140-6736(05)71048-5

2 Howson CP, Kinney M V, McDougall L, *et al.* Born Too Soon: Preterm birth matters. *Reprod Health* 2013;**10 Suppl 1**:S1. doi:10.1186/1742-4755-10-S1-S1

3 Blencowe H, Cousens S, Chou D, *et al.* Born Too Soon: The global epidemiology of 15 million preterm births. *Reprod Health* 2013;**10**:S2. doi:10.1186/1742-4755-10-S1-S2

4 Ananth C V, Vintzileos AM. Epidemiology of preterm birth and its clinical subtypes. *J Matern Fetal Neonatal Med* 2006;**19**:773–82. doi:10.1080/14767050600965882

5 Behrman R, ButlerAS, editors. *Institute of Medicine (IOM). Preterm Birth: Causes, Consequences, and Prevention.* Washington, D.C.: National Academies Press 2007. doi:10.17226/11622

6 Manuck TA, Esplin MS, Biggio J, *et al.* The phenotype of spontaneous preterm birth: application of a clinical phenotyping tool. *Am J Obstet Gynecol* 2015;**212**(4):487.e1-487.e11. doi:10.1016/j.ajog.2015.02.010

7 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**:422–33. doi:10.1097/GCO.0b013e328359823a

8 Conde-Agudelo A, Papageorghiou A, Kennedy S, *et al.* Novel biomarkers for the prediction of the spontaneous preterm birth phenotype: a systematic review and meta-analysis. *BJOG* 2011;**118**:1042–54. doi:10.1111/j.1471-0528.2011.02923.x

9 Hee L. Likelihood ratios for the prediction of preterm delivery with biomarkers. *Acta Obstet Gynecol Scand* 2011;**90**:1189–99. doi:10.1111/j.1600-0412.2011.01187.x

10 Goldenberg RL, Iams JD, Mercer BM, *et al.* The Preterm Prediction Study: toward a multiple-marker test for spontaneous preterm birth. *Am J Obstet Gynecol* 2001;**185**:643–51. doi:10.1067/mob.2001.116752

11 Iams JD, Goldenberg RL, Meis PJ, *et al.* The length of the cervix and the risk of spontaneous premature delivery. National Institute of Child Health and Human Development Maternal Fetal Medicine Unit Network. *N Engl J Med* 1996;**334**:567–72. doi:10.1056/NEJM199602293340904

12 Smith V, Devane D, Begley CM, *et al.* A systematic review and quality assessment of systematic reviews of fetal fibronectin and transvaginal length for predicting preterm birth. *Eur J Obstet Gynecol Reprod Biol* 2007;**133**:134–42. doi:10.1016/j.ejogrb.2007.03.005

13 Abbott DS, Hezelgrave NL, Seed PT, *et al.* Quantitative fetal fibronectin to predict preterm birth in asymptomatic women at high risk. *Obstet Gynecol* 2015;**125**:1168–

309   76. doi:10.1097/AOG.0000000000000754

310   14   Di Renzo GC. The great obstetrical syndromes. *J Matern neonatal Med* 2009;**22**:633–
311   5. doi:10.1080/14767050902866804

312   15   Brosens I, Pijnenborg R, Vercruysse L, *et al.* The 'Great Obstetrical Syndromes' are
313   associated with disorders of deep placentation. *Am J Obstet Gynecol* 2011;**204**:193–
314   201. doi:10.1016/j.ajog.2010.08.009

315   16   Practice bulletin no. 130: prediction and prevention of preterm birth. *Obstet Gynecol*
316   2012;**120**:964–73. doi:10.1097/AOG.0b013e3182723b1b

317   17   Goldenberg RL, Culhane JF, Iams JD, *et al.* Epidemiology and causes of preterm birth.
318   *Lancet* 2008;**371**:75–84. doi:10.1016/S0140-6736(08)60074-4

319   18   Honest H, Hyde CJ, Khan KS. Prediction of spontaneous preterm birth. *Curr Opin*
320   *Obstet Gynecol* 2012;**24**:422–33. doi:10.1097/GCO.0b013e328359823a

321   19   Horgan RP, Clancy OH, Myers JE, *et al.* An overview of proteomic and metabolomic
322   technologies and their application to pregnancy research. *BJOG* 2009;**116**:173–81.
323   doi:10.1111/j.1471-0528.2008.01997.x

324   20   Wishart DS, Feunang YD, Marcu A, *et al.* HMDB 4.0: the human metabolome database
325   for 2018. *Nucleic Acids Res* 2018;**46**:D608–17. doi:10.1093/nar/gkx1089

326   21   Dettmer K, Hammock BD. Metabolomics--a new exciting field within the omics
327   sciences. *Environ Health Perspect* 2004;**112**:A396-7.

328   22   Lucaroni F, Morciano L, Rizzo G, *et al.* Biomarkers for predicting spontaneous preterm
329   birth: an umbrella systematic review. *J Matern Neonatal Med* 2018;**31**:726–34.
330   doi:10.1080/14767058.2017.1297404

331   23   Romero R, Espinoza J, Gotsch F, *et al.* The use of high-dimensional biology (genomics,
332   transcriptomics, proteomics, and metabolomics) to understand the preterm
333   parturition syndrome. *BJOG* 2006;**113**:118–35. doi:10.1111/j.1471-
334   0528.2006.01150.x

335   24   Shamseer L, Moher D, Clarke M, *et al.* Preferred reporting items for systematic review
336   and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ*
337   2015;**350**:g7647. doi:10.1136/BMJ.G7647

338   25   Zhang A, Sun H, Wang P, *et al.* Modern analytical techniques in metabolomics
339   analysis. *Analyst* 2012;**137**:293–300. doi:10.1039/c1an15605e

340   26   Whiting PF, Rutjes AWS, Westwood ME, *et al.* QUADAS-2: A Revised Tool for the
341   Quality Assessment of Diagnostic Accuracy Studies. *Ann Intern Med* 2011;**155**:529.
342   doi:10.7326/0003-4819-155-8-201110180-00009

343   27   Rosselli D. The language of biomedical sciences. *Lancet* 2016;**387**:1720–1.
344   doi:10.1016/S0140-6736(16)30259-8

14

345

Author´s Contributions

347 RTS and RBFG will conduct the systematic review as independent first reviewers. JGC, RPJ

348 and DFBL will decide about conflicting decisions regarding papers selections. PNB, RPJ and

349 JGC participated in the systematic review conception, methodology and framework,

350 together will all the others co-authors.

Competing interests

360 All authors are carrying original research about metabolomics and presenting conferences

361 about this topic, including spontaneous preterm birth, preeclampsia, gestational diabetes

362 mellitus and fetal growth restriction. Philip N Baker is principal investigator of Metabolomics

363 Diagnostics Ltd, a company dedicated to develop innovative screening tests using

364 metabolomics technology.

Acknowledgements

15

366   Ana Paula de Morais e Oliveira, librarian of University of Campinas – Unicamp, Brazil, for

367   collaborating in developing search strategy and Rachel Hanisch for her suggestions to some

368   sections of the paper.

369   **Ethics approval and consent to participate**

370   This systematic review does not require ethical approval from the Research Council or Ethics

371   board.

16

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Searh strategy: #1 AND #2 AND #3

1 (OR for each term)

preterm birth
premature birth
premature infant
premature labor
extremely premature infant
premature obstetric labor
spontaneous preterm birth
extreme preterm birth
late preterm birth
moderate preterm birth
preterm premature rupture of membranes
preterm delivery
PROM
sPTB
preterm PROM
pPROM
p-PROM

2 (OR for each term)

metabolomic*
metabonomic*
metabolit*
lipidomic*
H NMR
proton NMR
proton nuclear magnetic resonance
liquid chromatogra*
UPLC
ultra-performance liquid chromatograph*
ultra performance liquid chromatograph*
HPLC
high performance liquid chromatograph*
high-performance liquid chromatograph*

3 (OR for each term)

pregnan*
antenat*
ante nat*
prenat*
pre nat*

PRISMA-P 2015 Checklist

This checklist has been adapted for use with systematic review protocol submissions to BioMed Central journals from Table 3 in Moher D et al**:** Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews* 2015 **4**:1

An Editorial from the Editors-in-Chief of *Systematic Reviews* details why this checklist was adapted - Moher D, Stewart L & Shekelle P**:** Implementing PRISMA-P: recommendations for prospective authors. *Systematic Reviews* 2016 **5**:15

Section/topic	#	Checklist item	Information reported		Line number(s)
			Yes	No	
ADMINISTRATIVE INFORMATION					
Title					
Identification	1a	Identify the report as a protocol of a systematic review	x	☐	2
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	☐	☐	n/a
Registration	2	If registered, provide the name of the registry (e.g., PROSPERO) and registration number in the Abstract	x	☐	48-49
Authors					
Contact	3a	Provide name, institutional affiliation, and e-mail address of all protocol authors; provide physical mailing address of corresponding author	x	☐	4-23
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review	x	☐	283-287
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments	☐	☐	n/a
Support					
Sources	5a	Indicate sources of financial or other support for the review	x	☐	275-282
Sponsor	5b	Provide name for the review funder and/or sponsor	x	☐	275-282
Role of sponsor/funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	x	☐	282
INTRODUCTION					
Rationale	6	Describe the rationale for the review in the context of what is already known	x	☐	51-87
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to	x	☐	94-95

1

2

Section/topic	#	Checklist item	Information reported		Line number(s)
			Yes	No	
		participants, interventions, comparators, and outcomes (PICO)			
METHODS					
Eligibility criteria	8	Specify the study characteristics (e.g., PICO, study design, setting, time frame) and report characteristics (e.g., years considered, language, publication status) to be used as criteria for eligibility for the review	x	☐	96-107
Information sources	9	Describe all intended information sources (e.g., electronic databases, contact with study authors, trial registers, or other grey literature sources) with planned dates of coverage	x	☐	119-125
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	x	☐	127-136
STUDY RECORDS					
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	x	☐	142-155
Selection process	11b	State the process that will be used for selecting studies (e.g., two independent reviewers) through each phase of the review (i.e., screening, eligibility, and inclusion in meta-analysis)	x	☐	157-161
Data collection process	11c	Describe planned method of extracting data from reports (e.g., piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	x	☐	163-167
Data items	12	List and define all variables for which data will be sought (e.g., PICO items, funding sources), any pre-planned data assumptions and simplifications	x	☐	142-155
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	x	☐	169-179
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	x	☐	189-196
DATA					
	15a	Describe criteria under which study data will be quantitatively synthesized	x	☐	198-209
Synthesis	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data, and methods of combining data from studies, including any planned exploration of consistency (e.g., I^2, Kendall's tau)	x	☐	185-187
	15c	Describe any proposed additional analyses (e.g., sensitivity or subgroup analyses, meta-regression)	x	☐	185-187

Section/topic	#	Checklist item	Information reported		Line number(s)
			Yes	No	
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	x	☐	198-202
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (e.g., publication bias across studies, selective reporting within studies)	☐	☐	n/a
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (e.g., GRADE)	☐	☐	n/a

BMJ Open

The use of metabolomics for predicting spontaneous preterm birth in asymptomatic pregnant women: protocol for a systematic review and meta-analysis

SCHOLARONE™
Manuscripts

1
2
3 **The use of metabolomics for predicting spontaneous preterm birth in asymptomatic**

4
5
6 **pregnant women: protocol for a systematic review and meta-analysis**

7
8
9

10
11 Renato T. Souza [1] (renatotsouzasp@gmail.com)
12
13 Rafael Bessa Freitas Galvão [1] (rafaelbfg@gmail.com)
14
15
16 Débora Farias Batista Leite [1,2] (deborafariasleite@gmail.com)
17
18 Renato Passini Jr [1] (passini@caism.unicamp.br)
19
20
21 Philip N. Baker [3] (philip.baker@leicester.ac.uk)
22
23 José Guilherme Cecatti [1] (cecatti@unicamp.br)
24
25
26 10

27
28 11 [1] Department of Obstetrics and Gynaecology, University of Campinas, Campinas, São Paulo,
29
30
31 12 Brazil.
32
33 13 [2] Clinics Hospital of Federal University of Pernambuco, Recife, Pernambuco, Brazil.
34
35 14 [3] College of Life Sciences, University of Leicester, England, United Kingdom.
36
37
38 15

39
40 16 This systematic review protocol is registered at PROSPERO platform (CRD42018100172). Available
41
42 17 from: http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42018100172.
43
44 18

45
46
47 19 Corresponding author:
48
49 20 José Guilherme Cecatti
50
51
52 21 101 Alexander Fleming, Cidade Universitária, campinas, São Paulo, Brazil.
53
54 22 ZIPCODE 13083-881
55
56
57 23 cecatti@unicamp.br
58
59 24 Word count: 2,233.
60

1

25 **ABSTRACT**

26 **Introduction**: Preterm birth (PTB) is the leading cause of neonatal mortality and short- and

27 long-term morbidity. The aetiology and pathophysiology of spontaneous PTB are still

28 unclear, which makes the identification of reliable and accurate predictor markers more

29 difficult, particularly for unscreened or asymptomatic women. Metabolomics biomarkers

30 have been demonstrated to be potentially accurate biomarkers for many disorders with

31 complex mechanisms such as PTB. Therefore, we aim to perform a systematic review of

32 metabolomics markers associated with spontaneous PTB. Our research question is "What is

33 the performance of metabolomics for predicting spontaneous preterm birth in

34 asymptomatic pregnant women?"

35 **Methods and analysis**: We will focus on studies assessing metabolomics techniques for

36 predicting spontaneous preterm birth in asymptomatic pregnant women. We will conduct a

37 comprehensive systematic review of the literature from the last 10 years. Only

38 observational cohort and case-control studies will be included. Our search strategy will be

39 carried out by two independent reviewers, who will scan title and abstract before carrying

40 out a full review of the article. The scientific databases to be explored include PubMed,

41 MedLine, ScieLo, EMBASE, LILACS, Web of Science, Scopus, and others.

42 **Ethics and dissemination:** This systematic review protocol does not require ethical approval.

43 We intend to disseminate our findings in scientific peer-reviewed journal, the Preterm

44 SAMBA study open access website, specialists' conferences, and to our funding agencies.

45 Registration details: This protocol is registered in PROSPERO platform (code

46 CRD42018100172).

47 **Keywords**: preterm birth, spontaneous preterm birth, metabolomics, biomarkers,

48 prediction, metabolome.

49 **Strengths and limitations of this study**

50 • This systematic review protocol takes into account some important aspects

51 regarding conducting a systematic review about spontaneous preterm birth and

52 metabolomics such as the criteria used for defining spontaneous preterm birth,

53 different population risk stratification, method used to estimate gestational age, and

54 metabolomics techniques details.

55 • Two independent reviewers are responsible for searching and selecting studies, as

56 also extracting data, and a third reviewer will resolve any disagreement.

57 • If possible, proper statistical methods will be applied to investigate metabolomics

58 accuracy in predicting spontaneous preterm birth.

59 • Possible limitations to this review include the different criteria applied for defining

60 spontaneous preterm birth, and the diverse population risk stratification.

3

62 **INTRODUCTION**

63 Spontaneous preterm birth (sPTB) is the leading cause of perinatal mortality and short- and

64 long-term morbidity [1,2]. It is defined as birth that occurs before 37 weeks gestation due to

65 spontaneous onset of labour or preterm premature rupture of membranes (pPROM) [3,4].

66 Several pathways and mechanisms linked with preterm birth have been proposed including,

67 neuroendocrine, vascular, immune-inflammatory, and behavioural processes [5]. More

68 specifically, several markers associated with uterine distension/contraction, decidual

69 inflammation/infection and activation of hypothalamic–pituitary–adrenal axis had been

70 studied in the past decades [5,6]. However, no single marker or combination of markers has

71 been found to be accurate enough for predicting sPTB [7–10]. History of previous preterm

72 birth, cervical length at second trimester and cervico-vaginal fetal fibronectin are the most

73 promising clinical tests for predicting spontaneous preterm, but they seem not to be

74 clinically useful for asymptomatic women. Sensitivity of short cervical length (<25mm) and

75 high cervico-vaginal fFN (>50ng/ml) are around 33-36% and 46%, respectively [11–13].

76 Preterm birth is a complex and multifactorial syndrome that possibly has a long pre-clinical

77 phase, maternal and fetal interactions, genetic and environmental influences, and adaptive

78 mechanisms [14,15]. These challenging aspects, and the presence of still unknown

79 underlying mechanisms, are the main limitations for the identification of an accurate

80 predictor for sPTB [16–18]. None of the predictors used in clinical practice, such as previous

81 history of preterm birth, infection (vaginal and urinary contaminants), fibronectin and

82 transvaginal ultrasonography cervical length demonstrated exceptional accuracy for

83 predicting spontaneous preterm birth [7]. An exploration of innovative approaches is

84 urgently required.

85 Metabolomics is the study of metabolites, through identification and quantification of low-

86 weight molecular particles, i.e. tens to hundreds thousands of intermediate products and

87 substrates of systems biology chemical reactions [19,20]. This novel approach has been

88 applied for identifying biomarkers and underlying biochemical pathways associated with

89 complex obstetrical syndromes as preeclampsia, fetal growth restriction, gestational

90 diabetes and preterm birth. In contrast to other "*Omics Sciences*" techniques, metabolomics

91 is more closely associated with the phenotype of the disease and might thus identify a more

92 robust and reliable set of predictors [21]. Importantly, implementing an adequate *Omics*

93 experimental design is crucial for metabolomics studies. Using different baseline population

94 (asymptomatic vs symptomatic or low- vs high-risk women for developing sPTB), study

95 designs (prospective cohorts, case-control or cross sectional studies), sources of samples

96 (amniotic fluid, vaginal fluid, blood, urine, hair, etc) and the timing of sample collection each

97 have significant effects on study findings and the consequent interpretation and

98 contribution to the current gap of knowledge [19].

99 Different reviews collating scientific knowledge regarding preterm birth

100 biomarkers/predictors has been conducted. Different methodology approaches has been

101 applied so far, including narrative, systematic and umbrella reviews, a more comprehensive

102 review that includes not only original studies but also other reviews [7,19,22,23]. At the best

103 of our knowledge, there is no systematic review on metabolomics markers. Therefore, we

104 aim to conduct a systematic review of original studies investigating the use of metabolomics

105 biomarkers for predicting spontaneous preterm birth in asymptomatic pregnant women.

106 This protocol describes the methods that will be applied in our systematic review.

107 **METHODS AND ANALYSIS**

108   The current systematic review proposal will be conducted, written and published following

109   the Preferred Reporting Items for Systematics Reviews and Meta-Analyses (PRISMA-P)

110   recommendations [24].   Also, it is properly registered at PROSPERO platform – code

111   CRD42018100172.

112   **Review question**

113   What is the performance of metabolomics for predicting spontaneous preterm birth in

114   asymptomatic pregnant women?

115   **Eligibility Criteria**

116   Original cohort or case-control studies involving asymptomatic pregnant women at the

117   moment of sample collection (exposure) and with samples analysed using metabolomics

118   techniques. Studies will be excluded if (1) they are cross-sectional studies, clinical trials,

119   editorials, letter to editors, case reports, expert opinions, commentaries, or any type of

120   review; (2) they describes only experimental studies with animals; or (3) they are duplicated

121   data (e.g. data published in conferences proceedings and, then, published again in scientific

122   journals). In this case, only the most complete publication will be considered, after

123   comparing and confirming that the same technique and metabolites were explored. Studies

124   published from 2008 to 2018 will be considered, and there will be no language restriction.

125   Before submitting this systematic review for publication, we will rerun the search strategy to

126   identify new studies that have been published after performing the first round of search.

127   **Participants**

128   The current review is interested in evaluating the performance of metabolomics biomarkers

129   for spontaneous preterm birth in asymptomatic pregnant women, which may contribute to

130   clinical practice, potentially providing information regarding onset of preterm labour.

131   Nevertheless, we aim to identify studies addressing only early predictors collected from

6

132 women who are in an early preclinical stage, which might contribute to a wider window of

133 opportunity for interventions and also to develop a widely reproducible screening test.

134 Asymptomatic pregnant women should not have regular uterine tightening/contractions or

135 signs of rupture of membranes (i.e. watery discharge). In addition, the study should

136 preferably have a standardized definition of spontaneous preterm birth, the outcome of

137 interest.

Information Sources

139 The search will be held in the following databases: PubMed, EMBASE, Scopus, CINAHL, and

140 Web of Science, BVS/BIREME, which includes the Latin American and Caribbean Health

141 Sciences Literature (LILACS), Medline and the Scientific Electronic Library Online (Scielo). In

142 addition, secondary sources of original studies will be explored such as Google Scholar,

143 hand-held searching of the reference list of eligible studies, conference proceedings, and

144 contact with authors when necessary.

Search Strategy

146 The following terms will be used in our search strategy for the different scientific databases:

147 (preterm birth, premature birth, premature infant, premature labor, extremely premature

148 infant, premature obstetric labor, spontaneous preterm birth, extreme preterm birth, late

149 preterm birth, moderate preterm birth, preterm premature rupture of membranes, preterm

150 delivery, PROM, sPTB, preterm PROM, pPROM, p-PROM) AND (metabolomic*,

151 metabonomic*, metabolit*, lipidomic*, H NMR, proton NMR, proton nuclear magnetic

152 resonance, liquid chromatogra*, gas chromatogra*, UPLC, ultra-performance liquid

153 chromatograph*, ultra-performance liquid chromatograph*, HPLC, high performance liquid

154 chrromatograph*, high-performance liquid chrromatograph*) AND (pregnan*, antenat*,

155 ante nat*, prenat*, pre nat*) (Supplementary Material). Respective adaptations in the

7

156 syntax of search for each database will be applied accordingly. No filters - such as "research

157 in animal's models" and "reviews" - will be used in our search strategy, as it will be excluded

158 according to eligibility criteria. The complete search strategy, including Boolean terms, is

159 provided as Supplementary Material.

Data Management

161 We will export search results to a reference manager (Mendeley®). Then, the following

162 information will be collected from each study using an appropriate form, which will be

163 entered in an Excel® spreadsheet: author's name, year of publication, country, study design,

164 number of participants with and without spontaneous preterm birth, type of metabolomics

165 analysis technique (liquid or gas chromatography, nuclear resonance), laboratory methods

166 for metabolites data acquiring (targeted or untargeted techniques, etc), subtype of preterm

167 birth (spontaneous preterm labour or pPROM), number of fetuses (singleton vs multiple),

168 gestational age when samples were collected, source of samples (type/site of tissue), low or

169 high-risk for preterm birth (authors criteria used to define the population will be collected)

170 and method applied to estimate gestational age. If possible, additional variables related to

171 spontaneous preterm birth categories (delivery before 28 weeks and before 34 weeks) will

172 be recorded for secondary analyses. Original authors will be contacted to clarify data, when

173 needed. Finally, we will check the biochemical class of identified metabolites in Human

174 Metabolome Database (HMDB, version 4.0) to explore and synthetize whether there are

175 common biological pathways associated with spontaneous preterm birth [20].

Selection Process

177 Two independent reviewers (RTS and RBFG) will be responsible for screening and selecting

178 studies initially according to title or abstract. Both researchers will read the full text of non-

179 excluded studies to discriminate eligibility. A third reviewer (DFBL) will consider any

180   disagreement; additional reviewers (RPJ, PNB and JGC) will be responsible for supervising all

181   steps and approving data extraction.

182   **Data Collection Process**

183   We will extract search results to a reference manager where all studies will be stored. Then,

184   included studies will be placed in a new folder. Finally, we will manually extract data of

185   interest from these included studies to an Excel® file. Each reviewer will have their own

186   reference manager account, file and folder and discrepant results will be discussed together

187   with the third reviewer.

188   **Outcomes and Prioritization**

189   The primary outcome is spontaneous preterm birth, defined as any birth occurred before 37

190   weeks of gestation due to spontaneous onset of labor or preterm premature rupture of

191   membranes. Secondary outcomes are:

192   1. Spontaneous preterm birth before 28 weeks;

193   2. Spontaneous preterm birth before 32 weeks;

194   3. Spontaneous preterm birth before 34 weeks;

195   The capacity to predict different degrees of sPTB (categories of gestational age) is important

196   as the extreme (<28wks), moderate (<32 weeks) and non-late preterm (<34wks) newborns

197   have different adverse outcomes compared to non-extreme (≥28wks); non-moderate

198   (≥32wks) or late (≥34 wks) preterm newborns.

199   Ideally, the method of gestational age estimation should be clearly reported. For instance, it

200   can be reported as estimated by last menstrual period (LMP) and confirmed by an early

201   ultrasound or only by an early ultrasound when LMP is unknown/uncertain.

202   **Index test**

9

203   Metabolomics techniques to predict spontaneous preterm birth is the diagnostic test of

204   interest. Metabolomics is a technique to identify and quantify metabolites from biological

205   samples using different type of platforms/equipment. The most common platforms include

206   gas, liquid chromatography or ultra-performance liquid chromatography coupled to a mass

207   spectrometer or a proton nuclear magnetic resonance [25]. The performance of the

208   different thresholds of each metabolite will be compared and summarized through

209   hierarchical summary receiver operator characteristic curve (HSROC) (meta-analysis)

210   according to the subgroups described above.

211   **Risk of Bias in individual Studies**

212   We will apply the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool [26]

213   to assess the risk of bias and applicability of primary diagnostic accuracy studies. Each study

214   will be classified as "low", "high" or "unclear" regarding risk of bias for each of the four

215   domains of QUADAS tool: Patient Selection, Index Test (metabolomics), Reference Standard

216   (occurrence of preterm birth), and Flow and Timing of participant's inclusion and follow-up.

217   For example, studies will be labelled as "low" risk of bias for Reference Standard when

218   definition of spontaneous preterm birth and gestational age estimation are clear; "high" risk

219   of bias would be considered when the moment of sample collection is not well described.

220   **Data Synthesis**

221   We will report details of identification, screening, eligibility and included studies using a flow

222   diagram, according to PRISMA recommendations [24]. Data from included studies will be

223   synthetized into tables according to the variables of interest. If possible, we will present

224   data meta-analysis according to study design, metabolomics technique and type of samples

225   analysed. We intend to perform subgroup analysis according to:

226   ●   Different metabolomics methods applied: gas or liquid chromatography coupled

10

227   with mass spectrometry or proton nuclear magnetic resonance;

228   ● Singleton and multiple pregnancies;

229   ● Low-risk and high-risk women for developing preterm birth;

230   ● Subtype of preterm birth: Spontaneous preterm birth exclusively due to

231   spontaneous onset of labour with intact membranes or sPTB due to premature

232   rupture of membranes.

233   ● Gestational age interval when samples were collected: 1st trimester, 2nd trimester

234   and 3rd trimester.

235   Heterogeneity will be assessed by Cochran's Q, Hotelling's T-squared (τ²) and I^2 tests. Funnel

236   plots and sensitivity and cumulative analyses will be applied for detection of temporal

237   trends and publication bias.

Potential anticipated limitations to this review

239   Firstly, although we have not considered any language restriction, we consider that there

240   might be a limitation in studies published entirely in non-English language. However, in the

241   last decade, more than 95% of scientific biomedical literature has been published in English

242   [27], then we consider this a minor selection bias. Secondly, we intend to stratify the groups

243   according to population risk. However, the characterization of low- or high-risk for

244   spontaneous preterm birth is controversial and lacks standardization, which might limit data

245   comparison and subgroup analysis. Finally, categorization of sPTB into spontaneous onset of

246   labour or pPROM is another topic of potential limitation - the recognition of the main initial

247   mechanism for preterm delivery might not always be possible. Even when specified, it might

248   provoke uncertainty and could limit further considerations regarding preterm phenotypes.

249   In addition, another limitation is that individual patient data will not be collected.

Patient and Public Involvement

11

251 Patients will not be directly involved in the study and no experience or direct impact from

252 their perspective can be discussed.

253

254 **ETHICS AND DISSEMINATION**

255 This systematic review does not require ethical approval from the Research Council or Ethics

256 board. We intend to disseminate our findings in scientific peer-reviewed journal, general

257 free access website of Preterm Screening and Metabolomics in Brazil and Auckland (Preterm

258 SAMBA) study, specialists' conferences, and to our funding agencies.

259 **DISCUSSION**

260 This systematic review will comprise current knowledge related with metabolomics in the

261 context of preterm birth prediction. Metabolomics science, a resourceful innovative field

262 that allows better understanding on pathophysiology of complex syndromes, may address

263 the main compounds associated with the spontaneous preterm delivery and, therefore,

264 motivate further researchers to validate early measurable predictors of preterm birth.

265 Metabolomics performance for predicting sPTB remains unclear and standardized and high-

266 quality studies are needed to clarify the clinical application of metabolites for predicting

267 sPTB. Nevertheless, metabolomics discovery studies commonly requires further validation

268 studies; reproducible methodology is crucial. This systematic review protocol will collate the

269 main potential early biomarkers, subgroup analysis and standardized definition for

270 spontaneous preterm birth to better understand metabolomics performance in predicting

271 sPTB and also to show its heterogeneity in terms of methodology (samples used,

272 metabolomics technique, definition of SPTB phenotype, etc). High performing predictors of

273 preterm birth will help combat this leading cause of neonatal mortality and morbidity.

1
2
3 274
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

13

276 **References**

277 1   Lawn JE, Cousens S, Zupan J. 4 million neonatal deaths: when? Where? Why? *Lancet*
278   2005;**365**:891–900. doi:10.1016/S0140-6736(05)71048-5

279 2   Howson CP, Kinney M V, McDougall L, *et al.* Born Too Soon: Preterm birth matters.
280   *Reprod Health* 2013;**10 Suppl 1**:S1. doi:10.1186/1742-4755-10-S1-S1

281 3   Blencowe H, Cousens S, Chou D, *et al.* Born Too Soon: The global epidemiology of 15
282   million preterm births. *Reprod Health* 2013;**10**:S2. doi:10.1186/1742-4755-10-S1-S2

283 4   Ananth C V, Vintzileos AM. Epidemiology of preterm birth and its clinical subtypes. *J*
284   *Matern Fetal Neonatal Med* 2006;**19**:773–82. doi:10.1080/14767050600965882

285 5   Behrman R, ButlerAS, editors. *Institute of Medicine (IOM). Preterm Birth: Causes,*
286   *Consequences, and Prevention.* Washington, D.C.: National Academies Press 2007.
287   doi:10.17226/11622

288 6   Manuck TA, Esplin MS, Biggio J, *et al.* The phenotype of spontaneous preterm birth:
289   application of a clinical phenotyping tool. *Am J Obstet Gynecol* 2015;**212**(4):487.e1-
290   487.e11. doi:10.1016/j.ajog.2015.02.010

291 7   Honest H, Hyde CJ, Khan KS. Prediction of spontaneous preterm birth: no good test
292   for predicting a spontaneous preterm birth. *Curr Opin Obstet Gynecol* 2012;**24**:422–
293   33. doi:10.1097/GCO.0b013e328359823a

294 8   Conde-Agudelo A, Papageorghiou A, Kennedy S, *et al.* Novel biomarkers for the
295   prediction of the spontaneous preterm birth phenotype: a systematic review and
296   meta-analysis. *BJOG* 2011;**118**:1042–54. doi:10.1111/j.1471-0528.2011.02923.x

297 9   Hee L. Likelihood ratios for the prediction of preterm delivery with biomarkers. *Acta*
298   *Obstet Gynecol Scand* 2011;**90**:1189–99. doi:10.1111/j.1600-0412.2011.01187.x

299 10 Goldenberg RL, Iams JD, Mercer BM, *et al.* The Preterm Prediction Study: toward a
300   multiple-marker test for spontaneous preterm birth. *Am J Obstet Gynecol*
301   2001;**185**:643–51. doi:10.1067/mob.2001.116752

302 11 Iams JD, Goldenberg RL, Meis PJ, *et al.* The length of the cervix and the risk of
303   spontaneous premature delivery. National Institute of Child Health and Human
304   Development Maternal Fetal Medicine Unit Network. *N Engl J Med* 1996;**334**:567–72.
305   doi:10.1056/NEJM199602293340904

306 12 Smith V, Devane D, Begley CM, *et al.* A systematic review and quality assessment of
307   systematic reviews of fetal fibronectin and transvaginal length for predicting preterm
308   birth. *Eur J Obstet Gynecol Reprod Biol* 2007;**133**:134–42.
309   doi:10.1016/j.ejogrb.2007.03.005

310 13 Abbott DS, Hezelgrave NL, Seed PT, *et al.* Quantitative fetal fibronectin to predict
311   preterm birth in asymptomatic women at high risk. *Obstet Gynecol* 2015;**125**:1168–

14

312   76. doi:10.1097/AOG.0000000000000754

313   14   Di Renzo GC. The great obstetrical syndromes. *J Matern neonatal Med* 2009;**22**:633–
314   5. doi:10.1080/14767050902866804

315   15   Brosens I, Pijnenborg R, Vercruysse L, *et al.* The 'Great Obstetrical Syndromes' are
316   associated with disorders of deep placentation. *Am J Obstet Gynecol* 2011;**204**:193–
317   201. doi:10.1016/j.ajog.2010.08.009

318   16   Practice bulletin no. 130: prediction and prevention of preterm birth. *Obstet Gynecol*
319   2012;**120**:964–73. doi:10.1097/AOG.0b013e3182723b1b

320   17   Goldenberg RL, Culhane JF, Iams JD, *et al.* Epidemiology and causes of preterm birth.
321   *Lancet* 2008;**371**:75–84. doi:10.1016/S0140-6736(08)60074-4

322   18   Honest H, Hyde CJ, Khan KS. Prediction of spontaneous preterm birth. *Curr Opin*
323   *Obstet Gynecol* 2012;**24**:422–33. doi:10.1097/GCO.0b013e328359823a

324   19   Horgan RP, Clancy OH, Myers JE, *et al.* An overview of proteomic and metabolomic
325   technologies and their application to pregnancy research. *BJOG* 2009;**116**:173–81.
326   doi:10.1111/j.1471-0528.2008.01997.x

327   20   Wishart DS, Feunang YD, Marcu A, *et al.* HMDB 4.0: the human metabolome database
328   for 2018. *Nucleic Acids Res* 2018;**46**:D608–17. doi:10.1093/nar/gkx1089

329   21   Dettmer K, Hammock BD. Metabolomics--a new exciting field within the omics
330   sciences. *Environ Health Perspect* 2004;**112**:A396-7.

331   22   Lucaroni F, Morciano L, Rizzo G, *et al.* Biomarkers for predicting spontaneous preterm
332   birth: an umbrella systematic review. *J Matern Neonatal Med* 2018;**31**:726–34.
333   doi:10.1080/14767058.2017.1297404

334   23   Romero R, Espinoza J, Gotsch F, *et al.* The use of high-dimensional biology (genomics,
335   transcriptomics, proteomics, and metabolomics) to understand the preterm
336   parturition syndrome. *BJOG* 2006;**113**:118–35. doi:10.1111/j.1471-
337   0528.2006.01150.x

338   24   Shamseer L, Moher D, Clarke M, *et al.* Preferred reporting items for systematic review
339   and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ*
340   2015;**350**:g7647. doi:10.1136/BMJ.G7647

341   25   Zhang A, Sun H, Wang P, *et al.* Modern analytical techniques in metabolomics
342   analysis. *Analyst* 2012;**137**:293–300. doi:10.1039/c1an15605e

343   26   Whiting PF, Rutjes AWS, Westwood ME, *et al.* QUADAS-2: A Revised Tool for the
344   Quality Assessment of Diagnostic Accuracy Studies. *Ann Intern Med* 2011;**155**:529.
345   doi:10.7326/0003-4819-155-8-201110180-00009

346   27   Rosselli D. The language of biomedical sciences. *Lancet* 2016;**387**:1720–1.
347   doi:10.1016/S0140-6736(16)30259-8

15

348

Author´s Contributions

350 RTS and RBFG will conduct the systematic review as independent first reviewers. JGC, RPJ

351 and DFBL will decide about conflicting decisions regarding papers selections. PNB, RPJ and

352 JGC participated in the systematic review conception, methodology and framework,

353 together will all the others co-authors.

Competing interests

363 All authors are carrying original research about metabolomics and presenting conferences

364 about this topic, including spontaneous preterm birth, preeclampsia, gestational diabetes

365 mellitus and fetal growth restriction. Philip N Baker is principal investigator of Metabolomics

366 Diagnostics Ltd, a company dedicated to develop innovative screening tests using

367 metabolomics technology.

Acknowledgements

16

369   Ana Paula de Morais e Oliveira, librarian of University of Campinas – Unicamp, Brazil, for

370   collaborating in developing search strategy and Rachel Hanisch for her suggestions to some

371   sections of the paper.

372   **Ethics approval and consent to participate**

373   This systematic review does not require ethical approval from the Research Council or Ethics

374   board.

17

Searh strategy: #1 AND #2 AND #3

1 (OR for each term)	preterm birth
	premature birth
	premature infant
	premature labor
	extremely premature infant
	premature obstetric labor
	spontaneous preterm birth
	extreme preterm birth
	late preterm birth
	moderate preterm birth
	preterm premature rupture of membranes
	preterm delivery
	PROM
	sPTB
	preterm PROM
	pPROM
	p-PROM
2 (OR for each term)	metabolomic*
	metabonomic*
	metabolit*
	lipidomic*
	H NMR
	proton NMR
	proton nuclear magnetic resonance
	liquid chromatogra*
	UPLC
	ultra-performance liquid chromatograph*
	ultra performance liquid chromatograph*
	HPLC
	high performance liquid chromatograph*
	high-performance liquid chromatograph*
3 (OR for each term)	pregnan*
	antenat*
	ante nat*
	prenat*
	pre nat*

1

PRISMA-P 2015 Checklist

This checklist has been adapted for use with systematic review protocol submissions to BioMed Central journals from Table 3 in Moher D et al**:** Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews* 2015 **4**:1

An Editorial from the Editors-in-Chief of *Systematic Reviews* details why this checklist was adapted - Moher D, Stewart L & Shekelle P**:** Implementing PRISMA-P: recommendations for prospective authors. *Systematic Reviews* 2016 **5**:15

Section/topic	#	Checklist item	Information reported Yes	No	Line number(s)
ADMINISTRATIVE INFORMATION					
Title					
Identification	1a	Identify the report as a protocol of a systematic review	x	☐	2
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	☐	☐	n/a
Registration	2	If registered, provide the name of the registry (e.g., PROSPERO) and registration number in the Abstract	x	☐	48-49
Authors					
Contact	3a	Provide name, institutional affiliation, and e-mail address of all protocol authors; provide physical mailing address of corresponding author	x	☐	4-23
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review	x	☐	283-287
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments	☐	☐	n/a
Support					
Sources	5a	Indicate sources of financial or other support for the review	x	☐	275-282
Sponsor	5b	Provide name for the review funder and/or sponsor	x	☐	275-282
Role of sponsor/funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	x	☐	282
INTRODUCTION					
Rationale	6	Describe the rationale for the review in the context of what is already known	x	☐	51-87
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to	x	☐	94-95

2

Section/topic	#	Checklist item	Information reported		Line number(s)
			Yes	**No**	
		participants, interventions, comparators, and outcomes (PICO)			
METHODS					
Eligibility criteria	8	Specify the study characteristics (e.g., PICO, study design, setting, time frame) and report characteristics (e.g., years considered, language, publication status) to be used as criteria for eligibility for the review	x	☐	96-107
Information sources	9	Describe all intended information sources (e.g., electronic databases, contact with study authors, trial registers, or other grey literature sources) with planned dates of coverage	x	☐	119-125
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	x	☐	127-136
STUDY RECORDS					
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	x	☐	142-155
Selection process	11b	State the process that will be used for selecting studies (e.g., two independent reviewers) through each phase of the review (i.e., screening, eligibility, and inclusion in meta-analysis)	x	☐	157-161
Data collection process	11c	Describe planned method of extracting data from reports (e.g., piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	x	☐	163-167
Data items	12	List and define all variables for which data will be sought (e.g., PICO items, funding sources), any pre-planned data assumptions and simplifications	x	☐	142-155
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	x	☐	169-179
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	x	☐	189-196
DATA					
Synthesis	15a	Describe criteria under which study data will be quantitatively synthesized	x	☐	198-209
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data, and methods of combining data from studies, including any planned exploration of consistency (e.g., I^2, Kendall's tau)	x	☐	185-187
	15c	Describe any proposed additional analyses (e.g., sensitivity or subgroup analyses, meta-regression)	x	☐	185-187

BMJ Open

3

Section/topic	#	Checklist item	Information reported		Line number(s)
			Yes	No	
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	x	☐	198-202
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (e.g., publication bias across studies, selective reporting within studies)	☐	☐	n/a
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (e.g., GRADE)	☐	☐	n/a

The use of metabolomics for predicting spontaneous preterm birth in asymptomatic pregnant women: protocol for a systematic review and meta-analysis

SCHOLARONE™
Manuscripts

The use of metabolomics for predicting spontaneous preterm birth in asymptomatic

pregnant women: protocol for a systematic review and meta-analysis

Renato T. Souza [1] (renatotsouzasp@gmail.com)

Rafael Bessa Freitas Galvão [1] (rafaelbfg@gmail.com)

Débora Farias Batista Leite [1,2] (deborafariasleite@gmail.com)

Renato Passini Jr [1] (passini@caism.unicamp.br)

Philip N. Baker [3] (philip.baker@leicester.ac.uk)

José Guilherme Cecatti [1] (cecatti@unicamp.br)

[1] Department of Obstetrics and Gynaecology, University of Campinas, Campinas, São Paulo,

Brazil.

[2] Clinics Hospital of Federal University of Pernambuco, Recife, Pernambuco, Brazil.

[3] College of Life Sciences, University of Leicester, England, United Kingdom.

Corresponding author:

José Guilherme Cecatti

101 Alexander Fleming, Cidade Universitária, campinas, São Paulo, Brazil.

ZIPCODE 13083-881

cecatti@unicamp.br

Word count: 2,233.

25 **ABSTRACT**

26 **Introduction**: Preterm birth (PTB) is the leading cause of neonatal mortality and short- and

27 long-term morbidity. The aetiology and pathophysiology of spontaneous PTB are still

28 unclear, which makes the identification of reliable and accurate predictor markers more

29 difficult, particularly for unscreened or asymptomatic women. Metabolomics biomarkers

30 have been demonstrated to be potentially accurate biomarkers for many disorders with

31 complex mechanisms such as PTB. Therefore, we aim to perform a systematic review of

32 metabolomics markers associated with spontaneous PTB. Our research question is "What is

33 the performance of metabolomics for predicting spontaneous preterm birth in

34 asymptomatic pregnant women?"

35 **Methods and analysis**: We will focus on studies assessing metabolomics techniques for

36 predicting spontaneous preterm birth in asymptomatic pregnant women. We will conduct a

37 comprehensive systematic review of the literature from the last 10 years. Only

38 observational cohort and case-control studies will be included. Our search strategy will be

39 carried out by two independent reviewers, who will scan title and abstract before carrying

40 out a full review of the article. The scientific databases to be explored include PubMed,

41 MedLine, ScieLo, EMBASE, LILACS, Web of Science, Scopus, and others.

42 **Ethics and dissemination:** This systematic review protocol does not require ethical approval.

43 We intend to disseminate our findings in scientific peer-reviewed journal, the Preterm

44 SAMBA study open access website, specialists' conferences, and to our funding agencies.

45 Registration details: This protocol is registered in PROSPERO platform (code

46 CRD42018100172).

47 **Keywords**: preterm birth, spontaneous preterm birth, metabolomics, biomarkers,

48 prediction, metabolome.

2

49 **Strengths and limitations of this study**

50 • This systematic review protocol takes into account some important aspects

51 regarding conducting a systematic review about spontaneous preterm birth and

52 metabolomics such as the criteria used for defining spontaneous preterm birth,

53 different population risk stratification, method used to estimate gestational age, and

54 metabolomics techniques details.

55 • Two independent reviewers are responsible for searching and selecting studies, as

56 also extracting data, and a third reviewer will resolve any disagreement.

57 • If possible, proper statistical methods will be applied to investigate metabolomics

58 accuracy in predicting spontaneous preterm birth.

59 • Possible limitations to this review include the different criteria applied for defining

60 spontaneous preterm birth, and the diverse population risk stratification.

3

62 **INTRODUCTION**

63 Spontaneous preterm birth (sPTB) is the leading cause of perinatal mortality and short- and

64 long-term morbidity [1,2]. It is defined as birth that occurs before 37 weeks gestation due to

65 spontaneous onset of labour or preterm premature rupture of membranes (pPROM) [3,4].

66 Several pathways and mechanisms linked with preterm birth have been proposed including,

67 neuroendocrine, vascular, immune-inflammatory, and behavioural processes [5]. More

68 specifically, several markers associated with uterine distension/contraction, decidual

69 inflammation/infection and activation of hypothalamic–pituitary–adrenal axis had been

70 studied in the past decades [5,6]. However, no single marker or combination of markers has

71 been found to be accurate enough for predicting sPTB [7–10]. History of previous preterm

72 birth, cervical length at second trimester and cervico-vaginal fetal fibronectin are the most

73 promising clinical tests for predicting spontaneous preterm, but they seem not to be

74 clinically useful for asymptomatic women. Sensitivity of short cervical length (<25mm) and

75 high cervico-vaginal fFN (>50ng/ml) are around 33-36% and 46%, respectively [11–13].

76 Preterm birth is a complex and multifactorial syndrome that possibly has a long pre-clinical

77 phase, maternal and fetal interactions, genetic and environmental influences, and adaptive

78 mechanisms [14,15]. These challenging aspects, and the presence of still unknown

79 underlying mechanisms, are the main limitations for the identification of an accurate

80 predictor for sPTB [16–18]. None of the predictors used in clinical practice, such as previous

81 history of preterm birth, infection (vaginal and urinary contaminants), fibronectin and

82 transvaginal ultrasonography cervical length demonstrated exceptional accuracy for

83 predicting spontaneous preterm birth [7]. An exploration of innovative approaches is

84 urgently required.

85 Metabolomics is the study of metabolites, through identification and quantification of low-

86 weight molecular particles, i.e. tens to hundreds thousands of intermediate products and

87 substrates of systems biology chemical reactions [19,20]. This novel approach has been

88 applied for identifying biomarkers and underlying biochemical pathways associated with

89 complex obstetrical syndromes as preeclampsia, fetal growth restriction, gestational

90 diabetes and preterm birth. In contrast to other "*Omics Sciences*" techniques, metabolomics

91 is more closely associated with the phenotype of the disease and might thus identify a more

92 robust and reliable set of predictors [21]. Importantly, implementing an adequate *Omics*

93 experimental design is crucial for metabolomics studies. Using different baseline population

94 (asymptomatic vs symptomatic or low- vs high-risk women for developing sPTB), study

95 designs (prospective cohorts, case-control or cross sectional studies), sources of samples

96 (amniotic fluid, vaginal fluid, blood, urine, hair, etc) and the timing of sample collection each

97 have significant effects on study findings and the consequent interpretation and

98 contribution to the current gap of knowledge [19].

99 Different reviews collating scientific knowledge regarding preterm birth

100 biomarkers/predictors has been conducted. Different methodology approaches has been

101 applied so far, including narrative, systematic and umbrella reviews, a more comprehensive

102 review that includes not only original studies but also other reviews [7,22–24]. At the best of

103 our knowledge, there is no systematic review on metabolomics markers. Therefore, we aim

104 to conduct a systematic review of original studies investigating the use of metabolomics

105 biomarkers for predicting spontaneous preterm birth in asymptomatic pregnant women.

106 This protocol describes the methods that will be applied in our systematic review.

107 **METHODS AND ANALYSIS**

108   The current systematic review proposal will be conducted, written and published following

109   the Preferred Reporting Items for Systematics Reviews and Meta-Analyses (PRISMA-P)

110   recommendations [25]. Also, it is properly registered at PROSPERO platform – code

111   CRD42018100172.

112   **Review question**

113   What is the performance of metabolomics for predicting spontaneous preterm birth in

114   asymptomatic pregnant women?

115   **Eligibility Criteria**

116   Original cohort or case-control studies involving asymptomatic pregnant women at the

117   moment of sample collection (exposure) and with samples analysed using metabolomics

118   techniques. Studies will be excluded if (1) they are cross-sectional studies, clinical trials,

119   editorials, letter to editors, case reports, expert opinions, commentaries, or any type of

120   review; (2) they describes only experimental studies with animals; or (3) they are duplicated

121   data (e.g. data published in conferences proceedings and, then, published again in scientific

122   journals). In this case, only the most complete publication will be considered, after

123   comparing and confirming that the same technique and metabolites were explored. Studies

124   published from 2008 to 2018 will be considered, and there will be no language restriction.

125   Before submitting this systematic review for publication, we will rerun the search strategy to

126   identify new studies that have been published after performing the first round of search.

127   **Participants**

128   The current review is interested in evaluating the performance of metabolomics biomarkers

129   for spontaneous preterm birth in asymptomatic pregnant women, which may contribute to

130   clinical practice, potentially providing information regarding onset of preterm labour.

131   Nevertheless, we aim to identify studies addressing only early predictors collected from

6

132 women who are in an early preclinical stage, which might contribute to a wider window of

133 opportunity for interventions and also to develop a widely reproducible screening test.

134 Asymptomatic pregnant women should not have regular uterine tightening/contractions or

135 signs of rupture of membranes (i.e. watery discharge). In addition, the study should

136 preferably have a standardized definition of spontaneous preterm birth, the outcome of

137 interest.

138 **Information Sources**

139 The search will be held in the following databases: PubMed, EMBASE, Scopus, CINAHL, and

140 Web of Science, BVS/BIREME, which includes the Latin American and Caribbean Health

141 Sciences Literature (LILACS), Medline and the Scientific Electronic Library Online (Scielo). In

142 addition, secondary sources of original studies will be explored such as Google Scholar,

143 hand-held searching of the reference list of eligible studies, conference proceedings, and

144 contact with authors when necessary.

145 **Search Strategy**

146 The following terms will be used in our search strategy for the different scientific databases:

147 (preterm birth, premature birth, premature infant, premature labor, extremely premature

148 infant, premature obstetric labor, spontaneous preterm birth, extreme preterm birth, late

149 preterm birth, moderate preterm birth, preterm premature rupture of membranes, preterm

150 delivery, PROM, sPTB, preterm PROM, pPROM, p-PROM) AND (metabolomic*,

151 metabonomic*, metabolit*, lipidomic*, H NMR, proton NMR, proton nuclear magnetic

152 resonance, liquid chromatogra*, gas chromatogra*, UPLC, ultra-performance liquid

153 chromatograph*, ultra-performance liquid chromatograph*, HPLC, high performance liquid

154 chrromatograph*, high-performance liquid chrromatograph*) AND (pregnan*, antenat*,

155 ante nat*, prenat*, pre nat*) (Supplementary Material). Respective adaptations in the

7

156 syntax of search for each database will be applied accordingly. No filters - such as "research

157 in animal's models" and "reviews" - will be used in our search strategy, as it will be excluded

158 according to eligibility criteria. The complete search strategy, including Boolean terms, is

159 provided as Supplementary Material.

160 **Data Management**

161 We will export search results to a reference manager (Mendeley®). Then, the following

162 information will be collected from each study using an appropriate form, which will be

163 entered in an Excel® spreadsheet: author's name, year of publication, country, study design,

164 number of participants with and without spontaneous preterm birth, type of metabolomics

165 analysis technique (liquid or gas chromatography, nuclear resonance), laboratory methods

166 for metabolites data acquiring (targeted or untargeted techniques, etc), subtype of preterm

167 birth (spontaneous preterm labour or pPROM), number of fetuses (singleton vs multiple),

168 gestational age when samples were collected, source of samples (type/site of tissue), low or

169 high-risk for preterm birth (authors criteria used to define the population will be collected)

170 and method applied to estimate gestational age. If possible, additional variables related to

171 spontaneous preterm birth categories (delivery before 28 weeks and before 34 weeks) will

172 be recorded for secondary analyses. Original authors will be contacted to clarify data, when

173 needed. Finally, we will check the biochemical class of identified metabolites in Human

174 Metabolome Database (HMDB, version 4.0) to explore and synthetize whether there are

175 common biological pathways associated with spontaneous preterm birth [20].

176 **Selection Process**

177 Two independent reviewers (RTS and RBFG) will be responsible for screening and selecting

178 studies initially according to title or abstract. Both researchers will read the full text of non-

179 excluded studies to discriminate eligibility. A third reviewer (DFBL) will consider any

180 disagreement; additional reviewers (RPJ, PNB and JGC) will be responsible for supervising all

181 steps and approving data extraction.

Data Collection Process

183 We will extract search results to a reference manager where all studies will be stored. Then,

184 included studies will be placed in a new folder. Finally, we will manually extract data of

185 interest from these included studies to an Excel® file. Each reviewer will have their own

186 reference manager account, file and folder and discrepant results will be discussed together

187 with the third reviewer.

Outcomes and Prioritization

189 The primary outcome is spontaneous preterm birth, defined as any birth occurred before 37

190 weeks of gestation due to spontaneous onset of labor or preterm premature rupture of

191 membranes. Secondary outcomes are:

    1. Spontaneous preterm birth before 28 weeks;

    2. Spontaneous preterm birth before 32 weeks;

    3. Spontaneous preterm birth before 34 weeks;

195 The capacity to predict different degrees of sPTB (categories of gestational age) is important

196 as the extreme (<28wks), moderate (<32 weeks) and non-late preterm (<34wks) newborns

197 have different adverse outcomes compared to non-extreme (≥28wks); non-moderate

198 (≥32wks) or late (≥34 wks) preterm newborns.

199 Ideally, the method of gestational age estimation should be clearly reported. For instance, it

200 can be reported as estimated by last menstrual period (LMP) and confirmed by an early

201 ultrasound or only by an early ultrasound when LMP is unknown/uncertain.

Index test

203   Metabolomics techniques to predict spontaneous preterm birth is the diagnostic test of

204   interest. Metabolomics is a technique to identify and quantify metabolites from biological

205   samples using different type of platforms/equipment. The most common platforms include

206   gas, liquid chromatography or ultra-performance liquid chromatography coupled to a mass

207   spectrometer or a proton nuclear magnetic resonance [26]. The performance of the

208   different thresholds of each metabolite will be compared and summarized through

209   hierarchical summary receiver operator characteristic curve (HSROC) (meta-analysis)

210   according to the subgroups described above. Considering that the raw data is not available

211   in the majority of the diagnostic test accuracy studies [27] and that metabolites levels are

212   usually reported as continuous variables, we intend to use a meta-analysis model based on

213   ROC curves [28]. Briefly, a two-parameter model, based on the estimation of α and β

214   parameters (using standard errors or maximum likelihood), will be applied as reported by

215   Kester & Buntinx [28]. Therefore, pooled ROC curves can be converted to a estimated ROC

216   curve with 95% confidence interval. This method can also be applied in categorical-ordinal

217   variables tests.

218   **Risk of Bias in individual Studies**

219   We will apply the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool [29]

220   to assess the risk of bias and applicability of primary diagnostic accuracy studies. Each study

221   will be classified as "low", "high" or "unclear" regarding risk of bias for each of the four

222   domains of QUADAS tool: Patient Selection, Index Test (metabolomics), Reference Standard

223   (occurrence of preterm birth), and Flow and Timing of participant's inclusion and follow-up.

224   For example, studies will be labelled as "low" risk of bias for Reference Standard when

225   definition of spontaneous preterm birth and gestational age estimation are clear; "high" risk

226   of bias would be considered when the moment of sample collection is not well described.

Data Synthesis

227

228 We will report details of identification, screening, eligibility and included studies using a flow

229 diagram, according to PRISMA recommendations [25]. Data from included studies will be

230 synthetized into tables according to the variables of interest. If possible, we will present

231 data meta-analysis according to study design, metabolomics technique and type of samples

232 analysed. We intend to perform subgroup analysis according to:

233 ● Different metabolomics methods applied: gas or liquid chromatography coupled

234 with mass spectrometry or proton nuclear magnetic resonance;

235 ● Singleton and multiple pregnancies;

236 ● Low-risk and high-risk women for developing preterm birth;

237 ● Subtype of preterm birth: Spontaneous preterm birth exclusively due to

238 spontaneous onset of labour with intact membranes or sPTB due to premature

239 rupture of membranes.

240 ● Gestational age interval when samples were collected: 1st trimester, 2nd trimester

241 and 3rd trimester.

242 Heterogeneity will be assessed by Cochran's Q, Hotelling's T-squared (τ^2) and I^2 tests. Funnel

243 plots and sensitivity and cumulative analyses will be applied for detection of temporal

244 trends and publication bias.

Potential anticipated limitations to this review

245

246 Firstly, although we have not considered any language restriction, we consider that there

247 might be a limitation in studies published entirely in non-English language. However, in the

248 last decade, more than 95% of scientific biomedical literature has been published in English

249 [30], then we consider this a minor selection bias. Secondly, we intend to stratify the groups

250 according to population risk. However, the characterization of low- or high-risk for

11

251 spontaneous preterm birth is controversial and lacks standardization, which might limit data

252 comparison and subgroup analysis. Finally, categorization of sPTB into spontaneous onset of

253 labour or pPROM is another topic of potential limitation - the recognition of the main initial

254 mechanism for preterm delivery might not always be possible. Even when specified, it might

255 provoke uncertainty and could limit further considerations regarding preterm phenotypes.

256 In addition, another limitation is that individual patient data will not be collected.

257 **Patient and Public Involvement**

258 Patients will not be directly involved in the study and no experience or direct impact from

259 their perspective can be discussed.

260

261 **ETHICS AND DISSEMINATION**

262 This systematic review does not require ethical approval from the Research Council or Ethics

263 board. We intend to disseminate our findings in scientific peer-reviewed journal, general

264 free access website of Preterm Screening and Metabolomics in Brazil and Auckland (Preterm

265 SAMBA) study, specialists' conferences, and to our funding agencies.

266 **DISCUSSION**

267 This systematic review will comprise current knowledge related with metabolomics in the

268 context of preterm birth prediction. Metabolomics science, a resourceful innovative field

269 that allows better understanding on pathophysiology of complex syndromes, may address

270 the main compounds associated with the spontaneous preterm delivery and, therefore,

271 motivate further researchers to validate early measurable predictors of preterm birth.

272 Metabolomics performance for predicting sPTB remains unclear and standardized and high-

273 quality studies are needed to clarify the clinical application of metabolites for predicting

12

274   sPTB. Nevertheless, metabolomics discovery studies commonly requires further validation

275   studies; reproducible methodology is crucial. This systematic review protocol will collate the

276   main potential early biomarkers, subgroup analysis and standardized definition for

277   spontaneous preterm birth to better understand metabolomics performance in predicting

278   sPTB and also to show its heterogeneity in terms of methodology (samples used,

279   metabolomics technique, definition of SPTB phenotype, etc). High performing predictors of

280   preterm birth will help combat this leading cause of neonatal mortality and morbidity.

281

13

283   **References**

284   1   Lawn JE, Cousens S, Zupan J. 4 million neonatal deaths: when? Where? Why? *Lancet*
285   2005;**365**:891–900. doi:10.1016/S0140-6736(05)71048-5

286   2   Howson CP, Kinney M V, McDougall L, *et al.* Born Too Soon: Preterm birth matters.
287   *Reprod Health* 2013;**10 Suppl 1**:S1. doi:10.1186/1742-4755-10-S1-S1

288   3   Blencowe H, Cousens S, Chou D, *et al.* Born Too Soon: The global epidemiology of 15
289   million preterm births. *Reprod Health* 2013;**10**:S2. doi:10.1186/1742-4755-10-S1-S2

290   4   Ananth C V, Vintzileos AM. Epidemiology of preterm birth and its clinical subtypes. *J
291   Matern Fetal Neonatal Med* 2006;**19**:773–82. doi:10.1080/14767050600965882

292   5   Behrman R, ButlerAS, editors. *Institute of Medicine (IOM). Preterm Birth: Causes,
293   Consequences, and Prevention.* Washington, D.C.: : National Academies Press 2007.
294   doi:10.17226/11622

295   6   Manuck TA, Esplin MS, Biggio J, *et al.* The phenotype of spontaneous preterm birth:
296   application of a clinical phenotyping tool. *Am J Obstet Gynecol* Published Online First:
297   February 2015. doi:10.1016/j.ajog.2015.02.010

298   7   Honest H, Hyde CJ, Khan KS. Prediction of spontaneous preterm birth: no good test
299   for predicting a spontaneous preterm birth. *Curr Opin Obstet Gynecol* 2012;**24**:422–
300   33. doi:10.1097/GCO.0b013e328359823a

301   8   Conde-Agudelo A, Papageorghiou A, Kennedy S, *et al.* Novel biomarkers for the
302   prediction of the spontaneous preterm birth phenotype: a systematic review and
303   meta-analysis. *BJOG An Int J Obstet Gynaecol* 2011;**118**:1042–54. doi:10.1111/j.1471-
304   0528.2011.02923.x

305   9   Hee L. Likelihood ratios for the prediction of preterm delivery with biomarkers. *Acta
306   Obstet Gynecol Scand* 2011;**90**:1189–99. doi:10.1111/j.1600-0412.2011.01187.x

307   10   Goldenberg RL, Iams JD, Mercer BM, *et al.* The Preterm Prediction Study: toward a
308   multiple-marker test for spontaneous preterm birth. *Am J Obstet Gynecol*
309   2001;**185**:643–51. doi:10.1067/mob.2001.116752

310   11   Iams JD, Goldenberg RL, Meis PJ, *et al.* The length of the cervix and the risk of
311   spontaneous premature delivery. National Institute of Child Health and Human
312   Development Maternal Fetal Medicine Unit Network. *N Engl J Med* 1996;**334**:567–72.
313   doi:10.1056/NEJM199602293340904

314   12   Smith V, Devane D, Begley CM, *et al.* A systematic review and quality assessment of
315   systematic reviews of fetal fibronectin and transvaginal length for predicting preterm
316   birth. *Eur J Obstet Gynecol Reprod Biol* 2007;**133**:134–42.
317   doi:10.1016/j.ejogrb.2007.03.005

318   13   Abbott DS, Hezelgrave NL, Seed PT, *et al.* Quantitative fetal fibronectin to predict

319   preterm birth in asymptomatic women at high risk. *Obstet Gynecol* 2015;**125**:1168–
320   76. doi:10.1097/AOG.0000000000000754

321   14   Di Renzo GC. The great obstetrical syndromes. *J Matern neonatal Med* 2009;**22**:633–
322   5. doi:10.1080/14767050902866804

323   15   Brosens I, Pijnenborg R, Vercruysse L, *et al.* The 'Great Obstetrical Syndromes' are
324   associated with disorders of deep placentation. *Am J Obstet Gynecol* 2011;**204**:193–
325   201. doi:10.1016/j.ajog.2010.08.009

326   16   Practice bulletin no. 130: prediction and prevention of preterm birth. *Obstet Gynecol*
327   2012;**120**:964–73. doi:10.1097/AOG.0b013e3182723b1b

328   17   Goldenberg RL, Culhane JF, Iams JD, *et al.* Epidemiology and causes of preterm birth.
329   *Lancet* 2008;**371**:75–84. doi:10.1016/S0140-6736(08)60074-4

330   18   Honest H, Hyde CJ, Khan KS. Prediction of spontaneous preterm birth. *Curr Opin*
331   *Obstet Gynecol* 2012;**24**:422–33. doi:10.1097/GCO.0b013e328359823a

332   19   Horgan RP, Clancy OH, Myers JE, *et al.* An overview of proteomic and metabolomic
333   technologies and their application to pregnancy research. *BJOG* 2009;**116**:173–81.
334   doi:10.1111/j.1471-0528.2008.01997.x

335   20   Wishart DS, Feunang YD, Marcu A, *et al.* HMDB 4.0: the human metabolome database
336   for 2018. *Nucleic Acids Res* 2018;**46**:D608–17. doi:10.1093/nar/gkx1089

337   21   Dettmer K, Hammock BD. Metabolomics--a new exciting field within the omics
338   sciences. *Environ Health Perspect* 2004;**112**:A396-
339   7.http://www.ncbi.nlm.nih.gov/pubmed/15159211 (accessed 5 Sep 2017).

340   22   Lucaroni F, Morciano L, Rizzo G, *et al.* Biomarkers for predicting spontaneous preterm
341   birth: an umbrella systematic review. *J Matern Neonatal Med* 2018;**31**:726–34.
342   doi:10.1080/14767058.2017.1297404

343   23   Horgan RP, Clancy OH, Myers JE, *et al.* An overview of proteomic and metabolomic
344   technologies and their application to pregnancy research. *BJOG* 2009;**116**:173–81.
345   doi:10.1111/j.1471-0528.2008.01997.x

346   24   Romero R, Espinoza J, Gotsch F, *et al.* The use of high-dimensional biology (genomics,
347   transcriptomics, proteomics, and metabolomics) to understand the preterm
348   parturition syndrome. *BJOG An Int J Obstet Gynaecol* 2006;**113**:118–35.
349   doi:10.1111/j.1471-0528.2006.01150.x

350   25   Shamseer L, Moher D, Clarke M, *et al.* Preferred reporting items for systematic review
351   and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ*
352   2015;**350**:g7647. doi:10.1136/BMJ.G7647

353   26   Zhang A, Sun H, Wang P, *et al.* Modern analytical techniques in metabolomics
354   analysis. *Analyst* 2012;**137**:293–300. doi:10.1039/c1an15605e

15

355   27   McGrath TA, Alabousi M, Skidmore B, *et al.* Recommendations for reporting of
356   systematic reviews and meta-analyses of diagnostic test accuracy: a systematic
357   review. *Syst Rev* 2017;**6**:194. doi:10.1186/s13643-017-0590-8

358   28   Kester ADM, Buntinx F. Meta-analysis of ROC Curves. *Med Decis Mak* 2000;**20**:430–9.
359   doi:10.1177/0272989X0002000407

360   29   Whiting PF, Rutjes AWS, Westwood ME, *et al.* QUADAS-2: A Revised Tool for the
361   Quality Assessment of Diagnostic Accuracy Studies. *Ann Intern Med* 2011;**155**:529.
362   doi:10.7326/0003-4819-155-8-201110180-00009

363   30   Rosselli D. The language of biomedical sciences. *Lancet* 2016;**387**:1720–1.
364   doi:10.1016/S0140-6736(16)30259-8

365

Author´s Contributions

367   RTS and RBFG will conduct the systematic review as independent first reviewers. JGC, RPJ

368   and DFBL will decide about conflicting decisions regarding papers selections. PNB, RPJ and

369   JGC participated in the systematic review conception, methodology and framework,

370   together will all the others co-authors.

Competing interests

380   All authors are carrying original research about metabolomics and presenting conferences

381   about this topic, including spontaneous preterm birth, preeclampsia, gestational diabetes

382   mellitus and fetal growth restriction. Philip N Baker is principal investigator of Metabolomics

383   Diagnostics Ltd, a company dedicated to develop innovative screening tests using

384   metabolomics technology.

385   **Acknowledgements**

386   Ana Paula de Morais e Oliveira, librarian of University of Campinas – Unicamp, Brazil, for

387   collaborating in developing search strategy and Rachel Hanisch for her suggestions to some

388   sections of the paper.

389   **Ethics approval and consent to participate**

390   This systematic review does not require ethical approval from the Research Council or Ethics

391   board.

17

Searrh strategy: #1 AND #2 AND #3

1 (OR for each term)	preterm birth
	premature birth
	premature infant
	premature labor
	extremely premature infant
	premature obstetric labor
	spontaneous preterm birth
	extreme preterm birth
	late preterm birth
	moderate preterm birth
	preterm premature rupture of membranes
	preterm delivery
	PROM
	sPTB
	preterm PROM
	pPROM
	p-PROM

2 (OR for each term)	metabolomic*
	metabonomic*
	metabolit*
	lipidomic*
	H NMR
	proton NMR
	proton nuclear magnetic resonance
	liquid chromatogra*
	UPLC
	ultra-performance liquid chromatograph*
	ultra performance liquid chromatograph*
	HPLC
	high performance liquid chromatograph*
	high-performance liquid chromatograph*

3 (OR for each term)	pregnan*
	antenat*
	ante nat*
	prenat*
	pre nat*

1

PRISMA-P 2015 Checklist

This checklist has been adapted for use with systematic review protocol submissions to BioMed Central journals from Table 3 in Moher D et al**:** Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews* 2015 **4**:1

An Editorial from the Editors-in-Chief of *Systematic Reviews* details why this checklist was adapted - Moher D, Stewart L & Shekelle P**:** Implementing PRISMA-P: recommendations for prospective authors. *Systematic Reviews* 2016 **5**:15

Section/topic	#	Checklist item	Information reported		Line number(s)
			Yes	No	
ADMINISTRATIVE INFORMATION					
Title					
Identification	1a	Identify the report as a protocol of a systematic review	x	☐	2
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	☐	☐	n/a
Registration	2	If registered, provide the name of the registry (e.g., PROSPERO) and registration number in the Abstract	x	☐	48-49
Authors					
Contact	3a	Provide name, institutional affiliation, and e-mail address of all protocol authors; provide physical mailing address of corresponding author	x	☐	4-23
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review	x	☐	283-287
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments	☐	☐	n/a
Support					
Sources	5a	Indicate sources of financial or other support for the review	x	☐	275-282
Sponsor	5b	Provide name for the review funder and/or sponsor	x	☐	275-282
Role of sponsor/funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	x	☐	282
INTRODUCTION					
Rationale	6	Describe the rationale for the review in the context of what is already known	x	☐	51-87
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to	x	☐	94-95

2

Section/topic	#	Checklist item	Information reported		Line number(s)
			Yes	**No**	
		participants, interventions, comparators, and outcomes (PICO)			
METHODS					
Eligibility criteria	8	Specify the study characteristics (e.g., PICO, study design, setting, time frame) and report characteristics (e.g., years considered, language, publication status) to be used as criteria for eligibility for the review	x	☐	96-107
Information sources	9	Describe all intended information sources (e.g., electronic databases, contact with study authors, trial registers, or other grey literature sources) with planned dates of coverage	x	☐	119-125
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	x	☐	127-136
STUDY RECORDS					
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	x	☐	142-155
Selection process	11b	State the process that will be used for selecting studies (e.g., two independent reviewers) through each phase of the review (i.e., screening, eligibility, and inclusion in meta-analysis)	x	☐	157-161
Data collection process	11c	Describe planned method of extracting data from reports (e.g., piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	x	☐	163-167
Data items	12	List and define all variables for which data will be sought (e.g., PICO items, funding sources), any pre-planned data assumptions and simplifications	x	☐	142-155
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	x	☐	169-179
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	x	☐	189-196
DATA					
Synthesis	15a	Describe criteria under which study data will be quantitatively synthesized	x	☐	198-209
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data, and methods of combining data from studies, including any planned exploration of consistency (e.g., I^2, Kendall's tau)	x	☐	185-187
	15c	Describe any proposed additional analyses (e.g., sensitivity or subgroup analyses, meta-regression)	x	☐	185-187

3

Section/topic	#	Checklist item	Information reported		Line number(s)
			Yes	No	
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	x	☐	198-202
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (e.g., publication bias across studies, selective reporting within studies)	☐	☐	n/a
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (e.g., GRADE)	☐	☐	n/a