

# BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email [info.bmjopen@bmj.com](mailto:info.bmjopen@bmj.com)

# BMJ Open

## Maternal metabolic profiling to assess fetal gestational age and predict preterm delivery

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2020-040647
Article Type:	Original research
Date Submitted by the Author:	18-May-2020
Complete List of Authors:	Sylvester, Karl; Stanford University Hao, Shiyong; Stanford University You, Jin; Stanford University Zheng, Le; Stanford University Tian, Lu; Stanford University Yao, Xiaoming; Hong Kong University Mo, Lihong; University of California San Francisco Fresno Center for Medical Education and Research Ladella, Subhashini; University of California San Francisco Fresno Center for Medical Education and Research Wong, Ronald; Stanford University Shaw, Gary M.; Stanford University Stevenson, David; Stanford University Cohen, Harvet; Stanford University Whitin, John; Stanford University McElhinney, Doff ; Stanford University Ling, Xuefeng; Stanford University,
Keywords:	Health informatics < BIOTECHNOLOGY & BIOINFORMATICS, Risk management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, OBSTETRICS

SCHOLARONE™  
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

1 **Maternal metabolic profiling to assess fetal gestational age and predict**  
2 **preterm delivery**

3 Karl G. SYLVESTER, MD<sup>1\*†</sup>, Shiyong HAO, PhD<sup>2,3\*</sup>, Jin YOU, PhD<sup>1\*</sup>, Le ZHENG,  
4 PhD<sup>2,3</sup>, Lu TIAN<sup>4</sup>, Xiaoming YAO, MD, PhD<sup>5</sup>, Lihong MO, MD<sup>6</sup>, Subhashini  
5 LADELLA, MD<sup>6</sup>, Ronald J. WONG, BA<sup>7</sup>, Gary M. SHAW, DrPh<sup>7</sup>, David K.  
6 STEVENSON, MD<sup>7</sup>, Harvey J. COHEN, MD, PhD<sup>7</sup>, John C. WHITIN, PhD<sup>7</sup>, Doff B.  
7 MCELHINNEY, MD<sup>2,3</sup>, Xuefeng B. LING, PhD<sup>1,3†</sup>

8  
9 <sup>1</sup>Department of Surgery, Stanford University School of Medicine, Stanford, CA, USA.

10 <sup>2</sup>Department of Cardiothoracic Surgery, Stanford University School of Medicine,  
11 Stanford, CA, USA.

12 <sup>3</sup>Clinical and Translational Research Program, Betty Irene Moore Children's Heart Center,  
13 Lucile Packard Children's Hospital, Palo Alto, CA, USA.

14 <sup>4</sup>Department of Health Research and Policy, Stanford University, Stanford, CA, USA.

15 <sup>5</sup>Translational Medicine Laboratory, Queen Mary Hospital, Hong Kong University, Hong  
16 Kong, China.

17 <sup>6</sup>Department of Obstetrics and Gynecology, University of California San Francisco-  
18 Fresno, Fresno, CA, USA.

19 <sup>7</sup>Department of Pediatrics, Stanford University School of Medicine, Stanford, CA, USA.

20  
21 †Correspondence: Xuefeng B. Ling, Department of Surgery, Stanford University,  
22 bxling@stanford.edu; Karl G. Sylvester, Department of Surgery, Stanford University,  
23 karls@stanford.edu

24 \*Co-first authors

25  
26 **Word count:** 280 for Abstract; 2973 for Main text.

1  
2  
3 27 **ABSTRACT**  
4  
5

6 28 **Objectives** The aim of this study was to develop a single blood test that could determine  
7  
8 29 gestational age and estimate the risk of preterm birth by measuring serum metabolites.  
9  
10 30 We hypothesized that serial metabolic modeling of serum analytes throughout pregnancy  
11  
12 31 could be used to describe fetal gestational age and project preterm birth with a high  
13  
14 32 degree of precision.  
15  
16  
17

18 33 **Study design** A retrospective cohort study  
19

20 34 **Setting** Two medical centers from US  
21  
22

23 35 **Participants** Thirty-six patients (20 full-term, 16 preterm) enrolled at Stanford  
24  
25 36 University were used to develop gestational age and preterm birth risk algorithms, 22  
26  
27 37 patients (9 full-term, 13 preterm) enrolled at the University of Alabama were used to  
28  
29 38 validate the algorithms.  
30  
31  
32

33 39 **Outcome measures** Maternal blood was collected serially throughout pregnancy.  
34  
35 40 Metabolic datasets were generated using mass spectrometry.  
36  
37

38 41 **Results** A model to determine gestational age was developed ( $R^2 = 0.98$ ) and validated  
39  
40 42 ( $R^2 = 0.81$ ). 66.7% of the estimates fell within  $\pm 1$  week of ultrasound results during  
41  
42 43 model validation. Significant disruptions from full-term pregnancy metabolic patterns  
43  
44 44 were observed in preterm pregnancies ( $R^2 = -0.68$ ). A separate algorithm to predict  
45  
46 45 preterm birth was developed utilizing a set of 10 metabolic pathways that resulted in an  
47  
48 46 area under the curve of 0.92 and a sensitivity of 0.86 during validation testing.  
49  
50  
51

52 47 **Conclusions** In this study metabolic profiling was used to develop and test a model for  
53  
54 48 determining gestational age during full-term pregnancy progression, and to determine  
55  
56  
57  
58  
59  
60

1  
2  
3 49 risk of preterm birth. With additional patient validation studies, these algorithms may be  
4  
5 50 used to identify at-risk pregnancies prompting alterations in clinical care, and to gain  
6  
7 51 biologic insights into the pathophysiology of preterm birth. Metabolic pathway-based  
8  
9 52 pregnancy modeling is a novel modality for investigation and clinical application  
10  
11 53 development.  
12  
13

14  
15 54 **Keywords:** Metabolic, gestational age, preterm birth, pathway  
16  
17

18 55  
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

1  
2  
3 **56 Strengths and limitations of this study**  
4  
5

- 6 57 • This study demonstrates a new non-invasive methodology for monitoring pregnancy  
7  
8 58 progression and identifying abnormal pregnancies at clinical settings.  
9  
10 59 • The insensitivity of the prediction model to gestational age (GA) window of sample  
11  
12 60 collection increases its flexibility and opportunity for potential clinical use.  
13  
14 61 • This study is among the first to propose a pathway-based computational methodology  
15  
16 62 to estimate GA and predict preterm birth.  
17  
18 63 • The overall cohort size is modest, and the distribution of sampling time are different  
19  
20 64 between patients and cohorts.  
21  
22 65 • It is a retrospective study; a larger prospective cohort study is necessary before  
23  
24 66 applying the estimates and prediction to a broader population for clinical utility.  
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

## 67 INTRODUCTION

68 Gestational age (GA) dating is a core element of standard prenatal care<sup>1-4</sup>. Prenatal  
69 ultrasound (US) is an established modality for estimating GA, monitoring fetal growth,  
70 and screening for fetal anomalies<sup>5</sup>. First trimester US imaging is the gold standard for  
71 GA determination, however there can be frequent discordance between US dating and a  
72 mother's last known menstrual period (LMP). In these cases, follow-up testing by US is  
73 utilized to more accurately estimate GA. US measurements are not currently used to  
74 determine risk of premature birth (PTB). The availability and expertise of US in  
75 disadvantaged areas is limited<sup>6</sup>. Therefore, there is a need to develop an alternative  
76 measure of fetal progression to estimate GA and pregnancy risk in a variety of settings  
77 and especially when US and LMP dates are unavailable or unreliable.

78 Compared with imaging methodologies, blood-based molecular testing may provide a  
79 more reproducible and precise modality in clinical applications for the frequent  
80 monitoring of health status and detection of early signs of disease. Genomic, gene  
81 expression, protein, and metabolite profiles measured in human blood have been  
82 increasingly utilized for the determination of disease risk and to gain disease specific  
83 pathophysiology insight. Attempts at estimating GA using molecular adaptations have  
84 included modeling of RNA, protein, or immune cell changes in maternal blood<sup>7-10</sup>, but  
85 not metabolites. Similarly, risk prediction of PTB in clinical settings is currently  
86 primarily based on maternal history. Biomarkers have been suggested from genetic and  
87 proteomic analyses, but less effort has been focused on understanding metabolic  
88 signatures of pregnancy<sup>11-16</sup>.



1  
2  
3 89 In this study, we hypothesized that longitudinal metabolic profiling of pregnancy reflects  
4  
5 90 the temporal progression of fetal development with a high degree of precision. Moreover,  
6  
7 91 we posited that if a normal pregnancy progression profile could be defined in metabolic  
8  
9 92 terms, then aberrations from the normal profile may identify a pregnancy at risk for PTB.  
10  
11 93 Herein, we have identified a panel of metabolic pathways measured in maternal serum  
12  
13 94 that provides an estimation of GA over the course of a full-term pregnancy. A second and  
14  
15 95 distinct set of metabolic pathways was also identified in maternal serum that could  
16  
17 96 distinguish pregnancies ending with PTB (< 35 weeks) from full-term ( $\geq$  37 weeks) with  
18  
19 97 a high degree of precision. The models were developed and validated using two  
20  
21 98 independent cohorts from two different institutions in order to test the robustness of the  
22  
23 99 biologic features driving the classifications. Our findings suggest that composite  
24  
25 100 metabolic panel modeling may serve as a reproducible and precision approach to GA  
26  
27 101 dating of pregnancy and prediction of PTB.  
28  
29  
30  
31  
32

## 33 102 **MATERIALS AND METHODS**

### 34 103 **Definition**

35  
36  
37  
38  
39 104 In this study, a full-term pregnancy was defined as a pregnancy ending with a delivery at  
40  
41 105  $\geq$  37 weeks. PTB was defined by delivery at < 35 weeks GA.  
42  
43

### 44 106 **Study design**

45  
46  
47 107 The study was conducted in two phases: (1) modeling to devise a metabolite-based  
48  
49 108 estimation of GA during full-term pregnancies; and (2) modeling to devise a metabolic  
50  
51 109 panel predictive of PTB (Fig. 1). In this study, the ‘gold’ standard of GA was US  
52  
53 110 measurement. Serum samples were collected in the 1<sup>st</sup>, 2<sup>nd</sup>, or 3<sup>rd</sup> trimester during  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 111 pregnancy for each individual woman. Each participant had 1 to 4 time-points collected  
4  
5 112 prior to delivery. Samples were provided by Stanford Hospital and Clinics (SU) and the  
6  
7 113 University of Alabama (UAB). Metabolic concentrations in each sample were measured  
8  
9 114 by targeted and untargeted mass spectrometry (MS) analysis. Models that estimated GA  
10  
11 115 or predicted PTB were developed using the SU cohort and validated using the UAB  
12  
13 116 cohort. The study was approved by the Institutional Review Board of both sites. All  
14  
15 117 samples were collected after informed consent was obtained. All statistical analyses were  
16  
17 118 done in R software.

### 119 **Targeted and global MS analysis**

120 Samples of full-term and preterm patients as well as quality control (QC) samples were  
121 injected into the MS. Targeted MS analysis was done through flow injection methods by  
122 using Ultimate 3000 Ultra-High-Performance Liquid Chromatography (UHPLC) system  
123 and Quantiva Triple Quadrupole Mass Spectrometer. Global (i.e. untargeted) MS analysis  
124 was done by using a Vanquish UHPLC system coupled to a Q Exactive plus mass  
125 spectrometer and Q Exactive HF hybrid quadrupole-Orbitrap mass spectrometer.

### 126 **Data preprocessing and metabolic identification**

127 A data pre-processing procedure was conducted to convert the raw data generated by MS  
128 analysis into a matrix of relative concentrations of metabolites versus samples<sup>17</sup>. This  
129 procedure was done by R package. Metabolic values in each sample were then  
130 normalized by the median values measured with QC samples to reduce the batch effects.  
131 Compounds detected by untargeted analyses were matched to metabolites in the Human  
132 Metabolome Database by putative identification<sup>18</sup>. Accurate mass was used for the

1  
2  
3 133 mapping. Metabolites were mapped to pathways using Kyoto Encyclopedia of Genes and  
4  
5 134 Genomes (KEGG) and Human Metabolome Database (HMDB). Only endogenous  
6  
7  
8 135 pathways were considered.  
9

### 10 136 **Metabolic compound selection, pathway computation, and model development**

11  
12  
13 137 Metabolites measured by targeted and untargeted MS were aggregated and filtered. The  
14  
15 138 remaining metabolites were mapped to pathways. The value of each pathway was  
16  
17  
18 139 calculated as the weighted sum of the normalized concentrations of metabolites on the  
19  
20 140 pathway divided by the number of metabolites. An XGBoost model was developed with  
21  
22 141 the pathway values of samples from full-term patients to estimate the GA. R-squared ( $R^2$ ;  
23  
24 142 goodness-of-fit of the model), root-mean-square error (RMSE), and error distribution  
25  
26  
27 143 were calculated to evaluate the model performance. A second XGBoost model was  
28  
29 144 developed to predict PTB. To evaluate the model performance, Mann–Whitney U tests  
30  
31 145 were used to compare the distribution of final predictive estimates, i.e., XGBoost model  
32  
33 146 values, on full-term and PTB samples. Results were compared with the insulin-like  
34  
35 147 growth factor-binding protein 4 (IBP4)/sex hormone-binding globulin (SHBG) signature  
36  
37 148 that is commercially available as a metabolic test for determining risk of PTB <sup>12</sup>.  
38  
39 149 Additional details of model development were described in Text A.1.  
40  
41  
42  
43

### 44 150 **Patient and Public Involvement statement**

45  
46  
47 151 This retrospective research was done without patient involvement. Patients were not  
48  
49 152 invited to comment on the study design and were not consulted to develop patient  
50  
51 153 relevant outcomes or interpret the results. Patients were not invited to contribute to the  
52  
53 154 writing or editing of this document for readability or accuracy.  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 **155 RESULTS**  
4

5  
6 **156 Samples**  
7

8  
9 157 As shown in Fig. 2, the SU cohort had 20 full-term pregnancies with 57 blood samples  
10  
11 158 (17, 32, and 8 collected in the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> trimesters, respectively) and 16 preterm  
12  
13 159 pregnancies with 32 blood samples (9, 19, and 4 collected in the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup>  
14  
15 160 trimesters, respectively). The UAB cohort had 9 full-term pregnancies with 13 blood  
16  
17 161 samples (8 and 5 in the 2<sup>nd</sup>, and 3<sup>rd</sup> trimesters, respectively) and 13 preterm pregnancies  
18  
19 162 with 22 blood samples (4 and 18 in the 1<sup>st</sup> and 2<sup>nd</sup> trimesters, respectively). In the SU  
20  
21 163 cohort, 2 (12.5%) were extremely preterm (< 28 weeks), and 5 (31.3%) were very  
22  
23 164 preterm (28–31 weeks). In the UAB cohort, 6 (46.2%) were extremely preterm, and 3  
24  
25 165 (23.1%) were very preterm. Demographics of the two cohorts are shown in Table 1.  
26  
27  
28  
29

30 **166 Table 1.** Maternal characteristics in SU and UAB cohorts  
31  
32 **167**

Characteristic	SU		<i>P</i>	UAB		<i>P</i>
	Full-term (n = 20)	Preterm (n = 16)		Full-term (n = 9)	Preterm (n = 13)	
Race, n (%)			<0.001			0.5
Asian	0	2 (12.5)		0	0	
White	20 (100)	5 (31.3)		0	2 (15.4)	
Black	0	1 (6.3)		9 (100)	10 (6.9)	
American Indian	0	1 (6.3)		0	0	
Pacific Islander	0	1 (6.3)		0	0	
Other/unknown	0	6 (37.5)		0	1 (7.7)	
Hispanic, n (%)	0	8 (50)	<0.001	0	1 (7.7)	0.9
Maternal Age, year, mean	31.9 (4.8)	29.8 (7.5)	0.3	25.6 (5.0)	27.5 (4.5)	0.4

(SD)

Gestational age at delivery, weeks, median (IQR)	39.5 (39,41)	32 (30,33)	<0.001	38 (37,39)	28 (26,32)	<0.001
Having previous pregnancy, n (%)	9 (45)	6 (37.5)	0.7	9 (100)	13 (100)	0.4
BMI, kg/m <sup>2</sup> , median (IQR)	22.3 (20.2,24.7)	27.6 (23.4,33.9)	0.003	30.4 (22.3,33.1)	26.5 (22.6,36.5)	0.8
History of PTB, n (%)	3 (15)	8 (50)	0.03	7 (77.8)	13 (100)	0.2

168

169

170 **LC-MS/MS metabolomics**

171 The study targeted 315 metabolites by LC-MS/MS, including 13 categories: acyl-  
 172 carnitine (11, 3.5%), amino acid (9, 2.9%), fatty acid (6, 1.9%), ceramide (12, 3.8%),  
 173 ceramide 1-phosphate (8, 2.5%), galactosylceramide (5, 1.6%), phosphatidyl acid (15,  
 174 4.8%), phosphatidylethanolamine (52, 16.5%), phosphatidylglycerol (5, 1.6%),  
 175 phosphatidylinositol (11, 3.5%), phosphatidylcholine (130, 41.3%), cholesteryl ester (16,  
 176 5.1%), and sphingomyelin (35, 11.1%). The study also identified 1627 positively-and 295  
 177 negatively-charged compounds through untargeted analyses. Together these formed the  
 178 initial set of 2237 compounds.

179 **Feature selection of GA estimation modeling**

180 Of the 2237 compounds, 115 had an absolute Pearson correlation coefficient of > 0.35  
 181 with GA. The cutoff of  $\pm 0.35$  was selected based on the false discovery rate (FDR)  
 182 values of the mapped pathways < 1% (Fig. A.1). The 115 compounds were mapped to 89  
 183 pathways, 33 of which were selected by the XGBoost model. The normalized value of

1  
2  
3 184 each pathway varied over the course of gestation (Fig. A.2). Univariate analysis of the 33  
4  
5 185 pathways is shown in Fig. A.3, and the top 10 pathways in the model is depicted in Fig. 3.  
6  
7  
8 186 The top 10 pathways included those associated in the metabolisms of:  
9  
10 187 glycerophospholipid, arginine and proline, thiamine, purine, butanoate, galactose, sulfur,  
11  
12 188 phenylalanine, and C5-branched dibasic acid.  
13  
14

### 15 189 **Performance of GA estimation**

16  
17  
18 190 The performance of GA estimates on full-term samples was similar in the development  
19  
20 191 phase (SU cohort,  $R^2 = 0.98$ , RMSE = 1.09) and the validation phase (UAB cohort,  $R^2 =$   
21  
22 192 0.81, RMSE = 2.36) (Fig. 4). In our validation testing, 66.7% of the estimates were  
23  
24 193 within  $\pm 1$  week of the US results (Fig. A.4).  
25  
26

27  
28 194 Intriguingly, model performance significantly deteriorated when applied to samples from  
29  
30 195 PTB pregnancies ( $R^2 = -0.68$  and RMSE = 6.6 in validation; see Fig. 4). It suggested that  
31  
32 196 the relationships between metabolic parameters and full-term pregnancies were not  
33  
34 197 maintained in PTB pregnancies. Furthermore, such disruptions were notable as early as  
35  
36 198 10 weeks' GA (Fig. 4) or early to mid-gestation. These findings prompted the  
37  
38 199 development of a metabolic-based model of PTB estimation.  
39  
40

### 41 200 **Performance of PTB prediction**

42  
43  
44  
45 201 Samples collected before 35 weeks' GA were used to develop a model that differentiated  
46  
47 202 PTB pregnancies from those full-term. As before, the model was developed with the SU  
48  
49 203 cohort that had 20 full-term (54 samples) and 16 preterm (32 samples) pregnancies, and  
50  
51 204 was validated with the UAB cohort that had 9 full-term (13 samples) and 13 preterm (22  
52  
53 205 samples) pregnancies. In total, 148 metabolic compounds (with Mann-Whitney U test  $P <$   
54  
55  
56  
57  
58  
59  
60

1  
2  
3 206 0.05) were mapped to 66 pathways (FDR < 1.5%; see Fig. A.5). Further model  
4  
5 207 development selected 10 pathways as strong predictors covering the metabolisms of  
6  
7 208 glycerophospholipid, sphingolipid, taurine and hypotaurine, arachidonic acid, secondary  
8  
9 209 bile acid biosynthesis, glycerolipid, cysteine and methionine, tryptophan, and arginine  
10  
11 210 and proline (Fig. 5).

12  
13  
14  
15 211 The level of prediction accuracy was maintained in the validation cohort ( $P = 5 \times 10^{-5}$ , area  
16  
17 212 under the curve [AUC] = 0.92; see Fig. 6). The prevalence-corrected positive predictive  
18  
19 213 values (PPVs) across model values (*i.e.* scores) were plotted based on the national PTB  
20  
21 214 prevalence in the United States (9.71%<sup>12 19</sup>; see Fig. A.6). A threshold value of 0.52 was  
22  
23 215 selected as a high-risk threshold for PTB, which was associated with a PPV of 0.61, a  
24  
25 216 relative risk (RR) of 6.3 compared to the United States population baseline (=

26  
27 217 0.61/9.71%), a sensitivity of 0.86 (19 of 22), and a specificity of 0.92 (12 of 13; Fig. 7).  
28  
29 218 The sensitivities and specificities with cutoff values are shown in Table A.1.  
30  
31  
32  
33

34 219 In the validation cohort, 12 of 13 full-term samples and 19 of 22 preterm samples were  
35  
36 220 classified correctly. The misclassified full-term sample was from a mother that delivered  
37  
38 221 at 37 weeks' GA. The 19 correctly classified PTB samples were from 13 PTB  
39  
40 222 pregnancies. Of the 13 pregnancies, 9 were identified as high risk at or earlier than 16  
41  
42 223 weeks' GA. The median gap between the time of identification and the delivery was 11  
43  
44 224 weeks' GA (IQR: 8, 15.5).

45  
46  
47  
48 225 To determine the performance of our metabolic model against existing models, a  
49  
50 226 comparison between the metabolic PTB risk model and the commercially available  
51  
52 227 IBP4/SHBG PTB test was performed and summarized in Text A.2.

53  
54  
55  
56 228 **Metabolite-based model and pathway-based model: a comparison**

1  
2  
3 229 To determine the effectiveness of model performance based upon robustness of biologic  
4  
5 230 features, we compared model performance using pathway or individual metabolite as  
6  
7  
8 231 selected features in estimating GA and predicting PTB. The performance of the pathway-  
9  
10 232 based models were significantly better than the metabolite-based models, with a lower  
11  
12 233 RMSE (Student's t-test  $P = 4 \times 10^{-3}$ ; Fig. A.7) and a larger AUC (DeLong test  $P = 0.03$ ;  
13  
14 234 Fig. A.8).

## 17 235 **DISCUSSION**

### 20 236 **Principal Findings**

23 237 In this study, metabolic modeling of maternal sera collected across gestation proved to be  
24  
25 238 a robust method of determining GA during pregnancy progression of term deliveries (>37  
26  
27 239 weeks' GA), in that it was validated in a population of women from a different center.  
28  
29  
30 240 Intriguingly, PTB pregnancies do not demonstrate the same temporal relationship as term  
31  
32 241 pregnancies upon metabolic modeling across gestation (Fig. 4). Indeed, PTB pregnancies  
33  
34 242 (<35 weeks' GA) demonstrate a marked departure from the term metabolic profile (Fig. 4)  
35  
36 243 that is not only dramatic ( $R^2 = 0.98$  train and  $0.81$  test for term model; compared to  $R^2 =$   
37  
38 244  $0.50$  train and  $-0.68$  test for PTB pregnancy in term model), but is also recognizable as  
39  
40 245 early as 10 weeks' GA as determined by the current standard of US dating. Recognizing  
41  
42 246 the metabolic pathway aberration of PTB pregnancies, a second model was developed  
43  
44 247 using metabolic pathway analyses to quantify the risk of PTB prior to 35 weeks' GA.  
45  
46 248 Once again, metabolic profiling proved to be robust in identifying PTB pregnancies with  
47  
48 249 a high degree of sensitivity (AUC 0.96 training; AUC 0.92 testing) and precision  
49  
50  
51 250 (training PPV 0.93 (0.78-0.99); testing PPV 0.95 (0.75-1). Taken together, this study  
52  
53  
54  
55  
56  
57  
58  
59  
60



1  
2  
3 251 demonstrated a powerful new methodology for monitoring pregnancy progression and  
4  
5 252 identifying abnormal pregnancies.  
6  
7

### 8 253 **Clinical and Research Implications**

9

10  
11 254 The potential clinical utility of developing a test for pregnancy monitoring is appealing.  
12  
13 255 There is a need to develop a more robust method than LMP and US that captures  
14  
15 256 pregnancy progression, a complex relationship of fetal and placental growth,  
16  
17 257 development, and function. To support these processes, there is a need for energy transfer  
18  
19 258 between mother and fetus throughout gestation. We therefore reasoned that metabolic  
20  
21 259 phenotyping would be ideally suited to capture this relationship. Despite a modest cohort  
22  
23 260 size, the results of metabolic modeling demonstrate a high degree of concordance with  
24  
25 261 clinical standard US dating performed by experts as reflected by 66.7% of model  
26  
27 262 estimates falling within  $\pm 1$  week of US results (Fig. A.4). Moreover, unlike the  
28  
29 263 deterioration experienced with US dating of pregnancy, metabolic modeling was shown  
30  
31 264 to achieve near equivalent performance in the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> trimesters, indicating the  
32  
33 265 potential for broad clinical applicability that might achieve independence of reliance on  
34  
35 266 accuracy of LMP or concordance among modality testing. The result of PTB prediction is  
36  
37 267 equally robust demonstrating a high degree of precision. Beyond relying on clinical  
38  
39 268 histories or self-reported symptoms, the model proposed here provides a molecular  
40  
41 269 classification that may be more accurate than current methods and further reflect a  
42  
43 270 comprehensive measure of aberrant pregnancy based on metabolic changes. In practice,  
44  
45 271 clinicians could use the PTB prediction model to differentiate high- from low-risk  
46  
47 272 patients. Low risk patients would then be subject to GA estimation panel testing, all from  
48  
49 273 the same blood draw.  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 274 A distinct advantage of the PTB risk prediction developed in this study is that it has a  
4  
5 275 wide window of sampling. Samples were collected broadly before 35 weeks' GA, which  
6  
7 276 is wider than the window of other well-established biomarkers such as fetal fibronectin  
8  
9  
10 277 (between 24 and 34 weeks' GA)<sup>13</sup>, IBP4/SHBG (19 to 21 weeks)<sup>12</sup>, and inter-alpha-  
11  
12 278 trypsin inhibitor heavy chain 4 protein (24 and 28 weeks)<sup>11</sup>. Relatively stable AUC  
13  
14 279 levels were maintained throughout the diagnostic window (Text A.2). The insensitivity of  
15  
16 280 the prediction model to GA at testing increases its flexibility and opportunity for potential  
17  
18 281 clinical use. An additional advantage of the model herein is the ability for early  
19  
20 282 identification of high-risk women. Although there is no standardized guideline for early-  
21  
22 283 gestation management of patients at risk of PTB delivery, metabolic modeling for PTB  
23  
24 284 risk may provide a not previously possible opportunity for early gestation risk mitigation.  
25  
26 285 Clinical trials have suggested that hormone treatment and maternal physical activity  
27  
28 286 modifications applied between 16 to 37 weeks' GA reduced the PTB rate of women who  
29  
30 287 were deemed at high risk due to a history of prior PTB delivery<sup>20 21</sup>. In many cases PTB  
31  
32 288 can not be prevented, however any opportunity is deemed highly desirable for even a  
33  
34 289 modest delay (1–2 weeks) in PTB or an enhanced ability to more accurately triage for  
35  
36 290 delivery to centers with the capability to manage profoundly premature neonates<sup>22-24</sup>.  
37  
38  
39  
40 291 This study is among the first to propose a pathway-based computational methodology to  
41  
42 292 estimate GA and predict PTB. Metabolic pathways are linked to chemical functions, and  
43  
44 293 the alteration or disruption of specific functions participate in disease phenotypes,  
45  
46 294 facilitating the use of pathways to function as higher-level biomarkers of diseases<sup>25</sup>. The  
47  
48 295 role of metabolic pathways in disease diagnosis has been explored in several preliminary  
49  
50 296 clinical studies<sup>26 27</sup>. Pathway performance in differentiating patients with disease from  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 297 healthy controls has been found to be effective compared to using individual metabolites  
4  
5 298 <sup>27</sup>. Similarly, we found the pathway-based models had less variability and higher  
6  
7  
8 299 sensitivity than metabolite-based models that were developed using the same population.  
9  
10 300 One plausible explanation for this observation may be attributed to the calculation of  
11  
12 301 pathway values, which represents the sum of individual metabolites and thus may  
13  
14 302 amplify association to outcome relationships. This hypothesis is supported by the FDR  
15  
16 303 comparison (Fig. A.7 and A.8): pathway-based analysis had lower FDR values than  
17  
18 304 metabolite models. This study adds to the exploration of the feasibility of using pathways  
19  
20  
21 305 for health monitoring and prediction.  
22

### 23 306 **Limitations**

24  
25  
26  
27 307 This study has several limitations. First, the overall cohort size was modest. Second,  
28  
29 308 blood samples were collected in a non-uniform manner with respect to GA timing and  
30  
31 309 time of day. The time between two adjacent samples corresponding to the same patient  
32  
33 310 varied. Third, the distribution of samples throughout pregnancy were different between  
34  
35 311 patients and cohorts. In the SU cohort, none of the full-term patients had samples  
36  
37 312 collected between 30 and 37 weeks. In the UAB cohort, none of the full-term patients had  
38  
39 313 sampling in the 1<sup>st</sup> trimester, and none of the PTB patients had sampling in the 3<sup>rd</sup>  
40  
41 314 trimester. Fourth, for methodologic reasons, not all serum analytes could be identified  
42  
43 315 and mapped to known metabolites. Fifth, the study was retrospective, and the participants  
44  
45 316 were solely from California and Alabama. A larger prospective cohort study is necessary  
46  
47 317 before applying the estimates and prediction to a broader population for clinical utility.  
48  
49  
50

### 51 318 **CONCLUSION**

52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 319 The present study demonstrates that maternal serum based metabolic profiling is a highly  
4  
5 320 sensitive and accurate method for determining GA and prediction of PTB. The pathway-  
6  
7 321 based analysis supports the hypothesis of the orderly metabolic progression of pregnancy  
8  
9 322 that can be reproducibly captured using metabolic profiling. The robustness of the  
10  
11 323 modeling reinforces the potential appeal for further clinical development and as a  
12  
13 324 platform to investigate the pathophysiology associated with aberrant fetal development  
14  
15 325 and pregnancy progression. This study is the first to report a single blood test for  
16  
17 326 metabolic pathway-based determination of GA dating, and early detection of PTB risk.  
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

1  
2  
3 328 **Acknowledgments:** The authors thank colleagues at the Stanford University Pediatric  
4  
5 329 Proteomics group and the March of Dimes Prematurity Research Center at Stanford  
6  
7  
8 330 University for critical discussions.  
9

10 331 **Funding:** This work was supported in part by the March of the Dimes Prematurity  
11  
12 332 Research Center at Stanford University, and Stanford Child Health Research Institute.  
13  
14  
15 333 The funders had no role in study design, data collection and analysis, decision to publish,  
16  
17 334 or preparation of the manuscript.  
18

19  
20 335 **Conflict of Interest:** The authors report no conflict of interest.  
21  
22

23 336 **Author contributions:** XBL, KGS, and HJC contributed to concept development and  
24  
25 337 design.  
26

27  
28 338 JY, RJW, and DKS contributed to the acquisition of data.  
29

30  
31 339 KGS, SH, LZ, XY, LT, LM, SL, RJW, GMS, DKS, JCW and DBM contributed to the  
32  
33 340 analysis and interpretation of data.  
34

35  
36 341 KGS and SH drafted the manuscript.  
37

38  
39 342 JY, LZ, LT, XY, LM, SL, RJW, GMS, DKS, HJC, JCW, DBM, and XBL critically  
40  
41 343 revised the manuscript.  
42

43  
44 344 All the authors gave final approval of the version to be submitted and agreed to be  
45  
46 345 accountable for all aspects of the work.  
47

48  
49 346 **Data and materials availability:** The datasets used and/or analyzed in this study are  
50  
51 347 available upon request to the corresponding author.  
52

53  
54 348  
55  
56  
57  
58  
59  
60

**References**

1. Brownfoot FC, Gagliardi DI, Bain E, et al. Different corticosteroids and regimens for accelerating fetal lung maturation for women at risk of preterm birth. *Cochrane Database Syst Rev* 2013(8):CD006764. doi: 10.1002/14651858.CD006764.pub3
2. Raju TN, Mercer BM, Burchfield DJ, et al. Periviable birth: executive summary of a Joint Workshop by the Eunice Kennedy Shriver National Institute of Child Health and Human Development, Society for Maternal-Fetal Medicine, American Academy of Pediatrics, and American College of Obstetricians and Gynecologists. *Journal of perinatology : official journal of the California Perinatal Association* 2014;34(5):333-42. doi: 10.1038/jp.2014.70
3. Vohr B. Long-term outcomes of moderately preterm, late preterm, and early term infants. *Clin Perinatol* 2013;40(4):739-51. doi: 10.1016/j.clp.2013.07.006
4. Pereira AP, Dias MA, Bastos MH, et al. Determining gestational age for public health care users in Brazil: comparison of methods and algorithm creation. *BMC Res Notes* 2013;6:60. doi: 10.1186/1756-0500-6-60
5. Peek MJ, Devonald KJ, Beilby R, et al. The value of routine early pregnancy ultrasound in the antenatal booking clinic. *Aust N Z J Obstet Gynaecol* 1994;34(2):140-3.
6. Jehan I, Zaidi S, Rizvi S, et al. Dating gestational age by last menstrual period, symphysis-fundal height, and ultrasound in urban Pakistan. *International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics* 2010;110(3):231-4. doi: 10.1016/j.ijgo.2010.03.030
7. Knight AK, Craig JM, Theda C, et al. An epigenetic clock for gestational age at birth based on blood methylation data. *Genome biology* 2016;17(1):206. doi: 10.1186/s13059-016-1068-z
8. Ngo TTM, Moufarrej MN, Rasmussen MH, et al. Noninvasive blood tests for fetal development predict gestational age and preterm delivery. *Science* 2018;360(6393):1133-36. doi: 10.1126/science.aar3819
9. Aghaeepour N, Lehallier B, Baca Q, et al. A proteomic clock of human pregnancy. *American journal of obstetrics and gynecology* 2018;218(3):347 e1-47 e14. doi: 10.1016/j.ajog.2017.12.208
10. Aghaeepour N, Ganio EA, McIlwain D, et al. An immune clock of human pregnancy. *Sci Immunol* 2017;2(15) doi: 10.1126/sciimmunol.aan2946
11. Esplin MS, Merrell K, Goldenberg R, et al. Proteomic identification of serum peptides predicting subsequent spontaneous preterm birth. *American journal of obstetrics and gynecology* 2011;204(5):391 e1-8. doi: 10.1016/j.ajog.2010.09.021
12. Saade GR, Boggess KA, Sullivan SA, et al. Development and validation of a spontaneous preterm delivery predictor in asymptomatic women. *American journal of obstetrics and gynecology* 2016;214(5):633 e1-33 e24. doi: 10.1016/j.ajog.2016.02.001
13. Peaceman AM, Andrews WW, Thorp JM, et al. Fetal fibronectin as a predictor of preterm birth in patients with symptoms: a multicenter trial. *American journal of obstetrics and gynecology* 1997;177(1):13-8.

- 1  
2  
3 395 14. Strauss JF, 3rd, Romero R, Gomez-Lopez N, et al. Spontaneous preterm birth:  
4 396 advances toward the discovery of genetic predisposition. *American journal of*  
5 397 *obstetrics and gynecology* 2018;218(3):294-314 e2. doi:  
6 398 10.1016/j.ajog.2017.12.009  
7  
8 399 15. Virgiliou C, Gika HG, Witting M, et al. Amniotic Fluid and Maternal Serum  
9 400 Metabolic Signatures in the Second Trimester Associated with Preterm  
10 401 Delivery. *Journal of proteome research* 2017;16(2):898-910. doi:  
11 402 10.1021/acs.jproteome.6b00845  
12  
13 403 16. Li J, Lu YP, Reichetzeder C, et al. Maternal PCaaC38:6 is Associated With Preterm  
14 404 Birth - a Risk Factor for Early and Late Adverse Outcome of the Offspring.  
15 405 *Kidney Blood Press Res* 2016;41(3):250-7. doi: 10.1159/000443428  
16 406 17. Dunn WB, Broadhurst D, Begley P, et al. Procedures for large-scale metabolic  
17 407 profiling of serum and plasma using gas chromatography and liquid  
18 408 chromatography coupled to mass spectrometry. *Nat Protoc* 2011;6(7):1060-  
19 409 83. doi: 10.1038/nprot.2011.335  
20  
21 410 18. Sumner LW, Amberg A, Barrett D, et al. Proposed minimum reporting standards  
22 411 for chemical analysis Chemical Analysis Working Group (CAWG)  
23 412 Metabolomics Standards Initiative (MSI). *Metabolomics* 2007;3(3):211-21.  
24 413 doi: 10.1007/s11306-007-0082-2  
25 414 19. Martin JA, Hamilton BE, Osterman MJ, et al. Births: final data for 2013. *Natl Vital*  
26 415 *Stat Rep* 2015;64(1):1-65.  
27  
28 416 20. Meis PJ, Klebanoff M, Thom E, et al. Prevention of recurrent preterm delivery by  
29 417 17 alpha-hydroxyprogesterone caproate. *The New England journal of*  
30 418 *medicine* 2003;348(24):2379-85. doi: 10.1056/NEJMoa035140  
31 419 21. Evenson KR, Siega-Riz AM, Savitz DA, et al. Vigorous leisure activity and  
32 420 pregnancy outcome. *Epidemiology* 2002;13(6):653-9. doi:  
33 421 10.1097/01.EDE.0000021463.45041.95  
34  
35 422 22. McIntire DD, Leveno KJ. Neonatal mortality and morbidity rates in late preterm  
36 423 births compared with births at term. *Obstetrics and gynecology*  
37 424 2008;111(1):35-41. doi: 10.1097/01.AOG.0000297311.33046.73  
38 425 23. Henderson-Smart DJ. The effect of gestational age on the incidence and duration  
39 426 of recurrent apnoea in newborn babies. *Aust Paediatr J* 1981;17(4):273-6.  
40  
41 427 24. Khashu M, Narayanan M, Bhargava S, et al. Perinatal outcomes associated with  
42 428 preterm birth at 33 to 36 weeks' gestation: a population-based cohort study.  
43 429 *Pediatrics* 2009;123(1):109-13. doi: 10.1542/peds.2007-3743  
44 430 25. Lee DS, Park J, Kay KA, et al. The implications of human metabolic network  
45 431 topology for disease comorbidity. *Proceedings of the National Academy of*  
46 432 *Sciences of the United States of America* 2008;105(29):9880-5. doi:  
47 433 10.1073/pnas.0802208105  
48  
49 434 26. Baumgartner C, Bohm C, Baumgartner D, et al. Supervised machine learning  
50 435 techniques for the classification of metabolic disorders in newborns.  
51 436 *Bioinformatics* 2004;20(17):2985-96. doi: 10.1093/bioinformatics/bth343  
52 437 27. Huang S, Chong N, Lewis NE, et al. Novel personalized pathway-based  
53 438 metabolomics models reveal key metabolic pathways for breast cancer  
54 439 diagnosis. *Genome medicine* 2016;8(1):34. doi: 10.1186/s13073-016-0289-9  
55  
56 440

## 441 **Figure Legends**

442 **Fig. 1.** Study design. Models were developed separately to estimate gestational age  
443 during full-term pregnancy, and to predict the risk of preterm birth. Both models were  
444 developed with the SU cohort and validated with the UAB cohort.

445 **Fig. 2.** Cohort construction. Each line represents an individual patient. Diamond and  
446 triangle markers indicate sample collection dates and delivery dates, respectively. The red  
447 dashed line represents 37 weeks' gestational age.

448 **Fig. 3.** The importance of the top 10 metabolic pathways in the gestational age estimation  
449 model. Pathways either positively or negatively correlated gestational age.

450 **Fig. 4.** Gestational age estimates of the gestational age model with the SU ( $R^2=0.98$ ,  
451  $RMSE=1.09$  weeks) and UAB cohorts ( $R^2=0.81$ ,  $RMSE=2.36$  weeks).

452 **Fig. 5.** (A) Univariate analysis of the 10 metabolic pathways in the preterm birth  
453 prediction model. Odds ratio of each pathway was calculated.  $*P<0.05$ ,  $**P<0.01$ ,  
454  $***P<0.005$ . (B) The importance of the metabolic pathways in the preterm birth  
455 prediction model. Pathways were either up- or down-regulated in relation to preterm birth.

456 **Fig. 6.** (A) Prediction of preterm birth risk grouped by full-term and preterm birth  
457 patients (top) and over the course of gestation (bottom). (B) AUC performance of the  
458 prediction in SU and UAB cohorts.  $P$  was calculated using Mann–Whitney U test. wks:  
459 weeks' gestational age.

460 **Fig. 7.** Performance of the preterm birth prediction model. (A) A contingency table  
461 showing the number of samples in each category. (B) Sensitivity, specificity, PPV, and  
462 NPV together with the 95% confidence intervals.



## 463 Appendix Captions

464 **Fig. A.1** False discovery rate (FDR) analysis of the metabolic pathways significantly  
465 associated with the GA in full-term pregnancies. Pearson  $|r|$  was calculated as the  
466 correlation between metabolite serological abundance and GA. Only the metabolites with  
467 a Pearson  $|r|$  higher than the threshold would be selected as part of the significant  
468 pathways. FDR was estimated by a permutation-based method (permutation  $N=1000$ ).

469 **Fig. A.2** Profile of the metabolic pathways in the GA estimation model over the course of  
470 gestation on SU cohort. All pathways are (A) positively or (B) negatively correlated to  
471 the GA (FDR<1%). Profile of each pathway was calculated as the weighted sum of the z-  
472 score normalized metabolite serological abundances divided by the number of  
473 metabolites. Means  $\pm$  standard errors at each time point were plotted.

474 **Fig. A.3** Univariate analysis of the 33 metabolic pathways in the GA estimation model.  
475 Pearson correlation coefficient of each pathway to GA was calculated. \* $P<0.05$ ,  
476 \*\* $P<0.01$ , \*\*\* $P<0.005$ .

477 **Fig. A.4** Comparison of GA estimates using the model and US measurements. (A)  
478 Distributions of differences between GA measured by US and GA estimated by the  
479 model, in T2 (weeks 14–27), T3 (weeks 28–40), and T2+T3.  $n$  represents the number of  
480 full-term patients included. (B) Error distribution of GA estimation on a combination of  
481 SU and UAB cohorts in T2, T3, and T2+T3.

482 **Fig. A.5** False discovery rate (FDR) analysis of the metabolic pathways significantly  
483 associated with PTB. Mann-Whitney U test  $P$  measured the difference in metabolite  
484 serological abundances between full-term pregnancies and pregnancies ending in PTB.

1  
2  
3 485 Only metabolites with a Mann-Whitney U test  $P$  lower than the threshold were selected  
4  
5 486 as part of the significant pathways. FDR was estimated by a permutation-based method  
6  
7 487 (permutation  $N=1000$ ).  
8  
9

10 488 **Fig. A.6** Stratification of patients by the classification model prediction on the UAB  
11  
12 489 cohort. PPV was corrected by bootstrapping the full-term patients to reach the population  
13  
14 490 PTB prevalence of 9.71% on singleton births. Two horizontal dashed lines represent the  
15  
16 491 population mean of PTB risk that is 9.71% (black) and the PPV (= 0.61; red) at the high-  
17  
18 492 risk cutoff. The grey dashed line indicates the high-risk cutoff value (= 0.52). The grey  
19  
20 493 area represents the 95% confidence interval of the PPV. The box plot at the bottom shows  
21  
22 494 the classification model value distribution stratified by the samples. GAB: GA at birth.  
23  
24 495 wks: weeks of gestation.  
25  
26  
27  
28

29 496 **Fig. A.7** (A) False discovery rate (FDR) analysis of the metabolites and metabolic  
30  
31 497 pathways significantly associated with GA in full-term pregnancies. Pearson  $|r|$  was  
32  
33 498 calculated as the correlation between metabolite serological abundance and GA. Only the  
34  
35 499 metabolites with a Pearson  $|r|$  higher than the threshold (=0.35) would be selected as part  
36  
37 500 of the significant pathways. FDR was estimated by a permutation-based method  
38  
39 501 (permutation  $N=1000$ ). (B) A comparison of RMSE of the GA estimation model trained  
40  
41 502 by pathways and the model trained by metabolites. All metabolites had a Pearson  $|r|>0.35$ .  
42  
43 503 RMSE was measured with the full-term samples of the validation (UAB) cohort.  
44  
45  
46  
47

48 504 **Fig. A.8** (A) False discovery rate (FDR) analysis of the metabolites and metabolic  
49  
50 505 pathways significantly associated with the PTB. Mann-Whitney U test  $P$  measured the  
51  
52 506 difference in metabolite serological abundances between full-term pregnancies and  
53  
54 507 pregnancies ending in PTB. Only the metabolites with a Mann-Whitney U test  $P$  lower  
55  
56  
57

1  
2  
3 508 than the threshold ( $=0.05$ ) would be selected as part of the significant pathways. FDR was  
4  
5 509 estimated by a permutation-based method (permutation  $N=1000$ ). (B) A comparison of  
6  
7 510 the AUC of the PTB classification model utilizing pathways and the model utilizing  
8  
9 511 metabolites. All the metabolites had a Mann-Whitney U test  $P < 0.05$ . AUC was  
10  
11 512 measured with the samples of the validation (UAB) cohort.  
12  
13  
14

15 513 **Table A.1** Sensitivity and specificity of the XGBoost model with respect to the cutoff  
16  
17 514 point.  
18  
19

20 515 **Text A.1** Metabolic compound selection, pathway computation, and model development  
21  
22

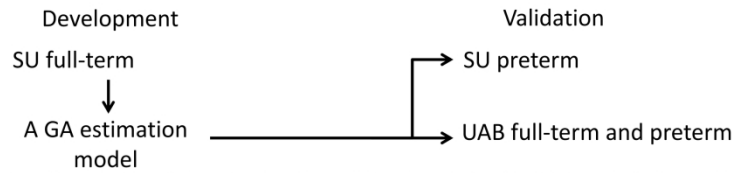
23 516 **Text A.2** Metabolite model vs. IBP4/SHBG in predicting PTB  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

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

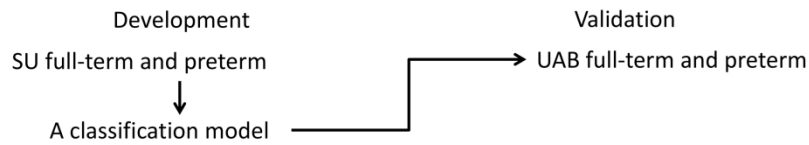
**A. Study cohort**

SU Cohort	UAB Cohort
20 full-term, 16 preterm	9 full-term, 13 preterm

**B. To estimate GA for full-term pregnancies**

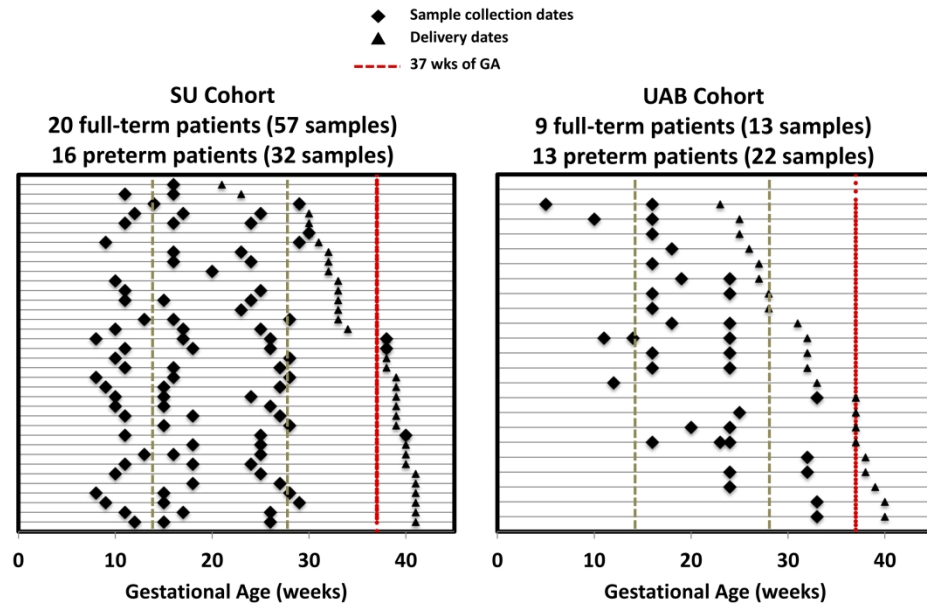


**C. To identify women at risk of PTB**



Study design. Models were developed separately to estimate gestational age during full-term pregnancy, and to predict the risk of preterm birth. Both models were developed with the SU cohort and validated with the UAB cohort.

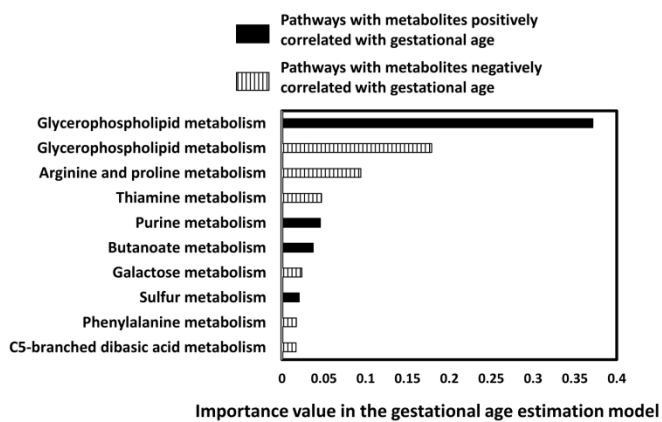
254x190mm (300 x 300 DPI)



31 Cohort construction. Each line represents an individual patient. Diamond and triangle markers indicate  
32 sample collection dates and delivery dates, respectively. The red dashed line represents 37 weeks'  
33 gestational age.

34 254x190mm (300 x 300 DPI)

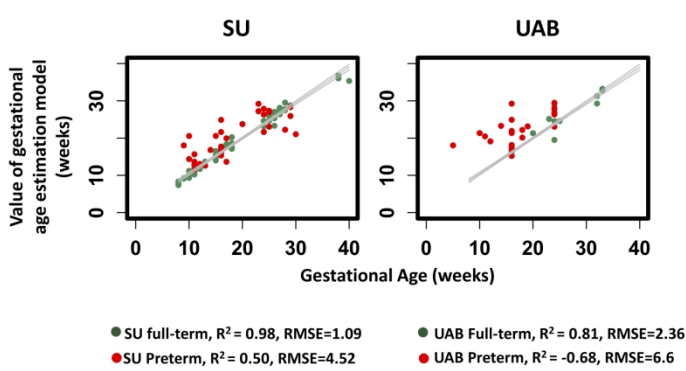
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



The importance of the top 10 metabolic pathways in the gestational age estimation model. Pathways either positively or negatively correlated gestational age.

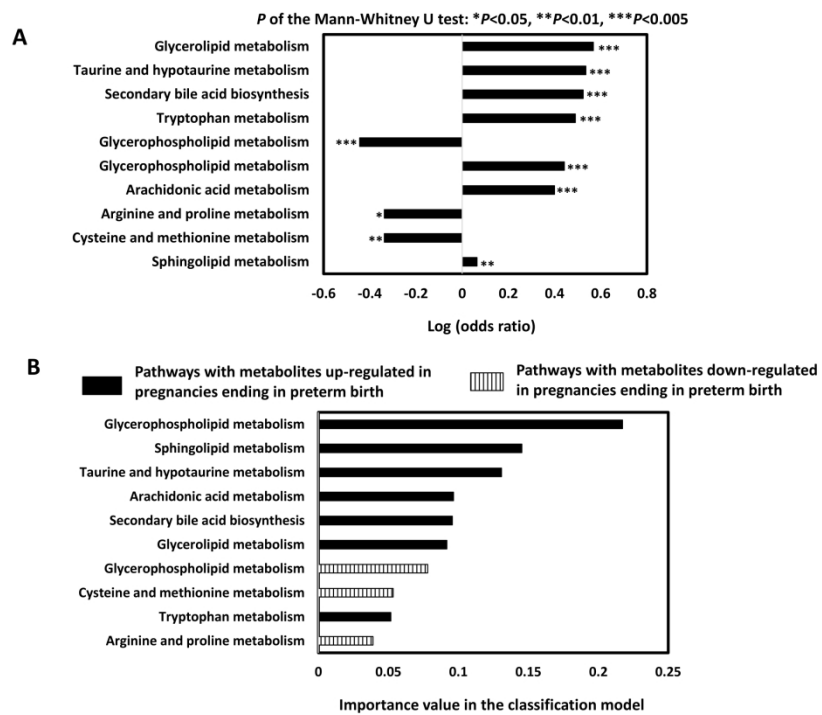
254x190mm (300 x 300 DPI)

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



Gestational age estimates of the gestational age model with the SU (R2=0.98, RMSE=1.09 weeks) and UAB cohorts (R2 = 0.81, RMSE = 2.36 weeks).

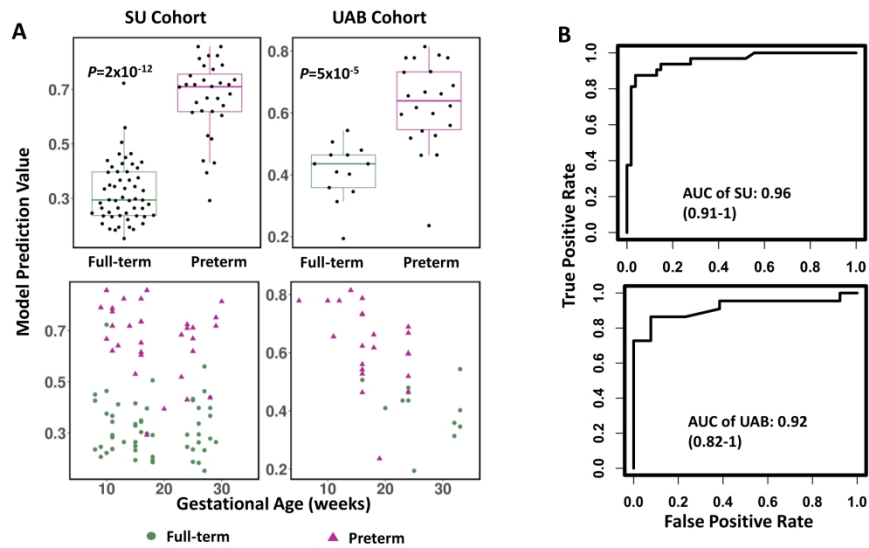
254x190mm (300 x 300 DPI)



(A) Univariate analysis of the 10 metabolic pathways in the preterm birth prediction model. Odds ratio of each pathway was calculated. \**P*<0.05, \*\**P*<0.01, \*\*\**P*<0.005. (B) The importance of the metabolic pathways in the preterm birth prediction model. Pathways were either up- or down-regulated in relation to preterm birth.

254x190mm (300 x 300 DPI)





(A) Prediction of preterm birth risk grouped by full-term and preterm birth patients (top) and over the course of gestation (bottom). (B) AUC performance of the prediction in SU and UAB cohorts. P was calculated using Mann-Whitney U test. wks: weeks' gestational age.

254x190mm (300 x 300 DPI)

**A**

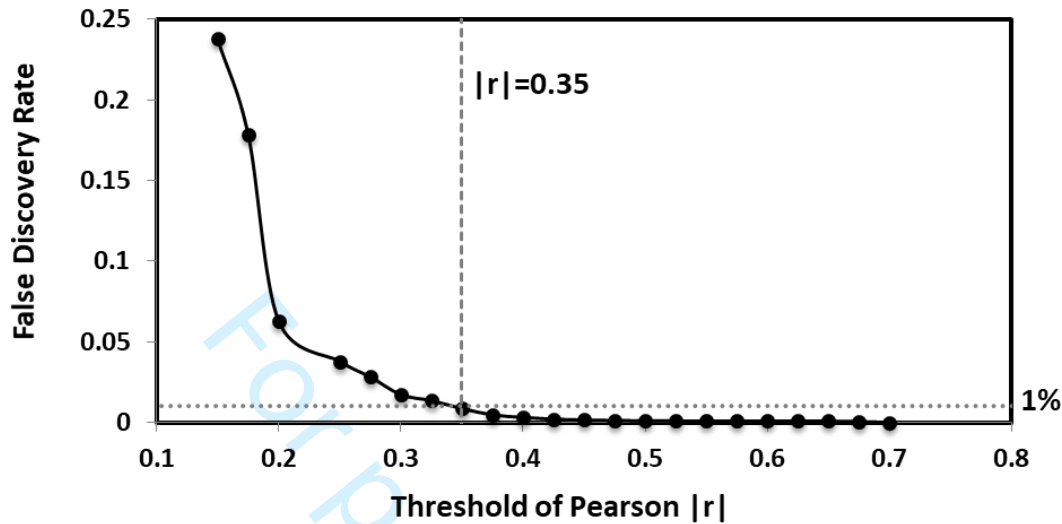
	<b>SU</b>			<b>UAB</b>	
	Preterm	Full-term		Preterm	Full-term
Classified as Preterm	28	2	Classified as Preterm	19	1
Classified as Full-term	4	52	Classified as Full-term	3	12

**B**

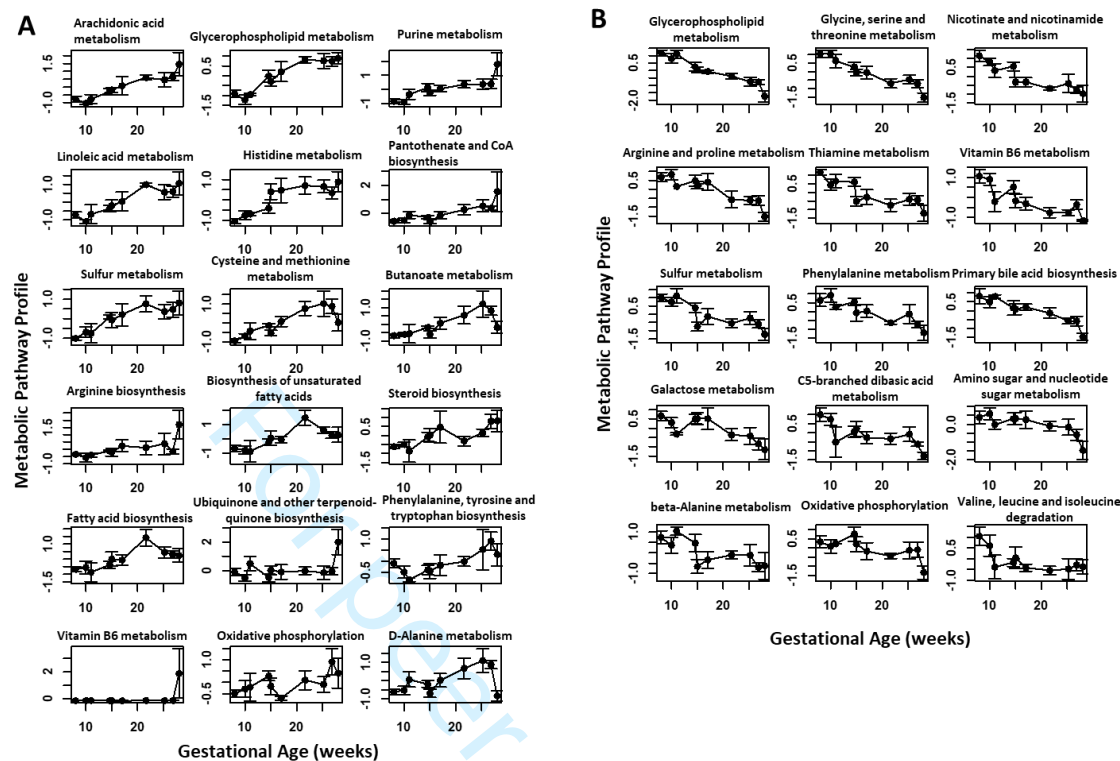
Cohort	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
<b>SU</b>	0.88 (0.71-0.97)	0.96 (0.87-1)	0.93 (0.78-0.99)	0.93 (0.83-0.98)
<b>UAB</b>	0.86 (0.65-0.97)	0.92 (0.64-1)	0.95 (0.75-1)	0.80 (0.52-0.96)

Performance of the preterm birth prediction model. (A) A contingency table showing the number of samples in each category. (B) Sensitivity, specificity, PPV, and NPV together with the 95% confidence intervals.

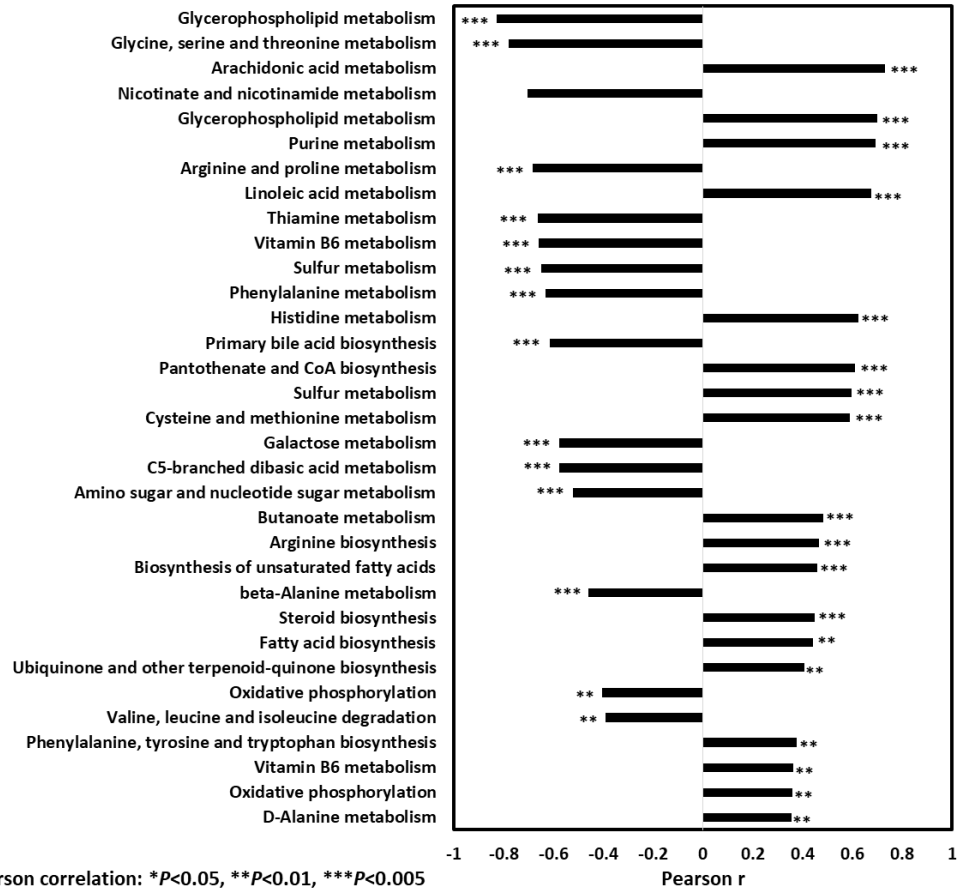
254x190mm (300 x 300 DPI)



**Fig. A.1.** False discovery rate (FDR) analysis of the metabolic pathways significantly associated with the GA in full-term pregnancies. Pearson  $|r|$  was calculated as the correlation between metabolite serological abundance and GA. Only the metabolites with a Pearson  $|r|$  higher than the threshold would be selected as part of the significant pathways. FDR was estimated by a permutation-based method (permutation  $N=1000$ ).



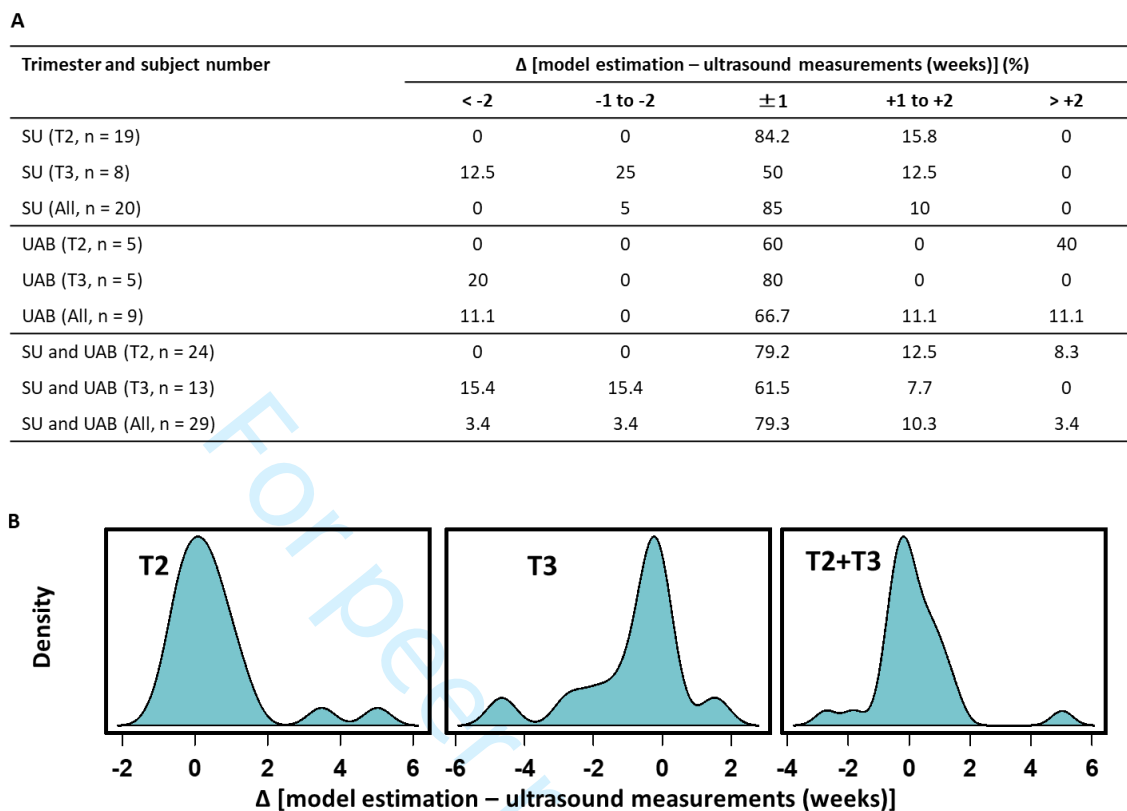
**Fig. A.2.** Profile of the metabolic pathways in the GA estimation model over the course of gestation on SU cohort. All pathways are (A) positively or (B) negatively correlated to the GA (FDR<1%). Profile of each pathway was calculated as the weighted sum of the z-score normalized metabolite serological abundances divided by the number of metabolites. Mean  $\pm$  standard error of the mean at each time point was plotted.



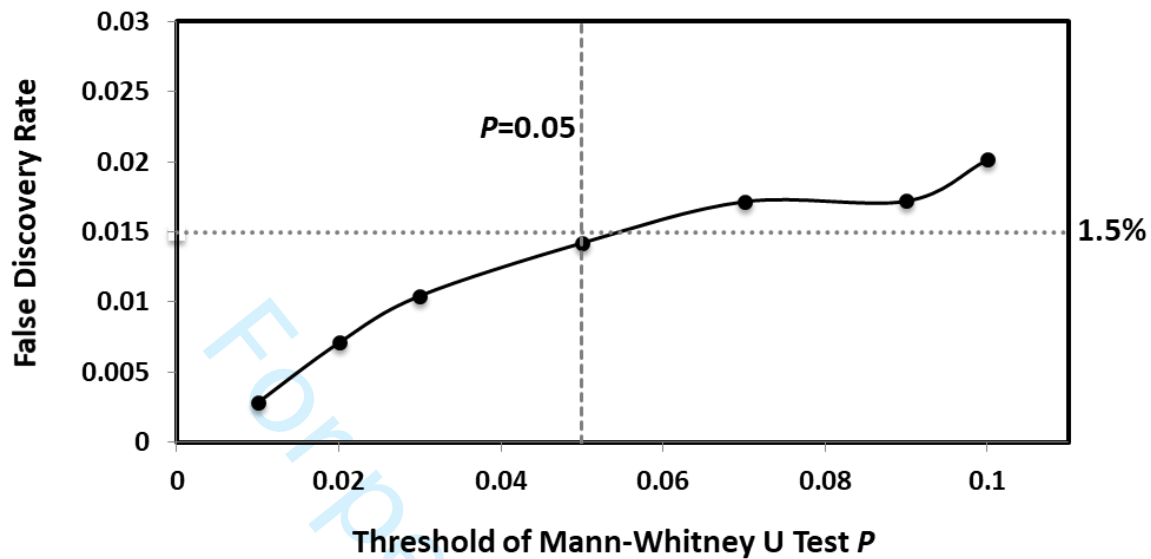
**Fig. A.3.** Univariate analysis of the 33 metabolic pathways in the GA estimation model.

Pearson correlation coefficient  $r$  of each pathway to GA was calculated. \* $P<0.05$ ,

\*\* $P<0.01$ , \*\*\* $P<0.005$ .

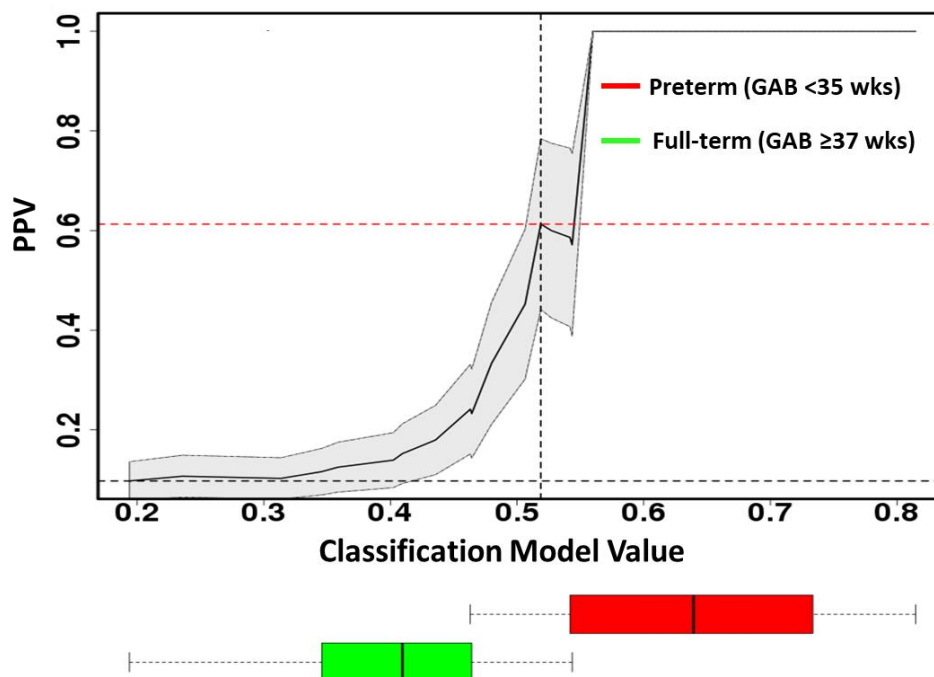


**Fig. A.4.** Comparison of GA estimates using the model and US measurements. (A) Distributions of differences between GA measured by US and GA estimated by the model, in T2 (weeks 14–27), T3 (weeks 28–40), and T2+T3. n represents the number of full-term patients included. (B) Error distribution of GA estimation on a combination of SU and UAB cohorts in T2, T3, and T2+T3.



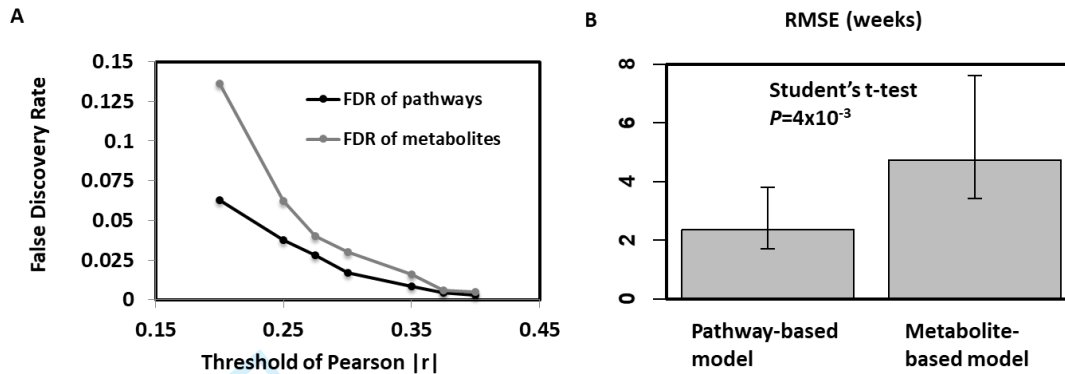
**Fig. A.5.** False discovery rate (FDR) analysis of the metabolic pathways significantly associated with PTB. Mann-Whitney U test  $P$  measured the difference in metabolite serological abundances between full-term pregnancies and pregnancies ending in PTB. Only metabolites with a Mann-Whitney U test  $P$  lower than the threshold were selected as part of the significant pathways. FDR was estimated by a permutation-based method (permutation  $N=1000$ ).

Population-corrected PPV: 0.61. which is **6.3** times higher than the general population in the US (9.71%)

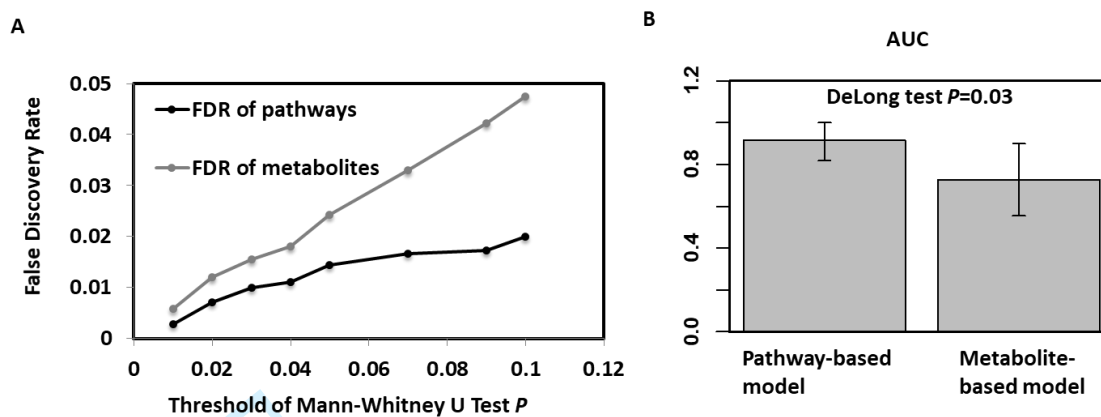


**Fig. A.6.** Stratification of patients by the classification model prediction on the UAB cohort. PPV was corrected by bootstrapping the full-term patients to reach the population PTB prevalence of 9.71% on singleton births. Two horizontal dashed lines represent the population mean of PTB risk that is 9.71% (black) and the PPV (= 0.61; red) at the high-risk cutoff. The grey dashed line indicates the high-risk cutoff value (= 0.52). The grey area represents the 95% confidence interval of the PPV. The box plot at the bottom shows the classification model value distribution stratified by the samples. GAB: gestational age at birth. wks: weeks' GA.





**Fig. A.7.** (A) False discovery rate (FDR) analysis of the metabolites and metabolic pathways significantly associated with the GA in full-term pregnancies. Pearson  $|r|$  was calculated as the correlation between metabolite serological abundance and GA. Only the metabolites with a Pearson  $|r|$  higher than the threshold ( $=0.35$ ) would be selected as part of the significant pathways. FDR was estimated by a permutation-based method (permutation  $N=1000$ ). (B) A comparison of RMSE of the GA estimation model trained by pathways and the model trained by metabolites. All metabolites had a Pearson  $|r| > 0.35$ . RMSE was measured with the full-term samples of the validation (UAB) cohort.



**Fig. A.8.** (A) False discovery rate (FDR) analysis of the metabolites and metabolic pathways significantly associated with the PTB. Mann-Whitney U test  $P$  measured the difference in metabolite serological abundances between full-term pregnancies and pregnancies ending in PTB. Only the metabolites with a Mann-Whitney U test  $P$  lower than the threshold ( $=0.05$ ) would be selected as part of the significant pathways. FDR was estimated by a permutation-based method (permutation  $N=1000$ ). (B) A comparison of the AUC of the preterm birth classification model utilizing pathways and the model utilizing metabolites. All the metabolites had a Mann-Whitney U test  $P < 0.05$ . AUC was measured with the samples of the validation (UAB) cohort.

**Table A.1.** Sensitivity and specificity of the XGBoost model with respect to the cutoff point.

Cutoff	Cohort	Sensitivity	Specificity	Number of preterm samples identified by the model
<b>0.4</b>	<b>SU</b>	0.94	0.78	30
	<b>UAB</b>	0.95	0.31	21
<b>0.5</b>	<b>SU</b>	0.88	0.94	28
	<b>UAB</b>	0.86	0.85	19
<b>0.6</b>	<b>SU</b>	0.81	0.98	26
	<b>UAB</b>	0.59	1	13
<b>0.7</b>	<b>SU</b>	0.53	0.98	17
	<b>UAB</b>	0.32	1	7

## **Text A.1 Metabolic compound selection, pathway computation, and model development**

### *GA estimation*

Metabolites measured by targeted and untargeted MS were aggregated and filtered using Pearson correlation coefficient analyses in relation to GA. The remaining metabolites were mapped to pathways. The value of each pathway was calculated as the weighted sum of the normalized concentrations of metabolites on the pathway divided by the number of metabolites. The weight of each metabolite was the absolute value of the Pearson correlation coefficient in relation to GA. Metabolites having positive or negative coefficients were aggregated separately. That is, a pathway could have two values, one for metabolites positively correlated to GA, and the other for those negatively correlated to GA.

A supervised, cross-validated machine-learning technique XGBoost was developed with the pathway values of samples from full-term patients in the SU cohort. An ensemble of regression trees was generated to give a score estimating the GA. The model was validated on the UAB cohort. For a patient that had multiple samples, an ‘integrated’ GA estimate was calculated by shifting the GA estimates of every sample to a reference point for obtaining the median. Error distribution of GA estimation based on patients was calculated as the distribution of the differences between the ‘integrated’ GA estimates and the US measurement.

### *PTB prediction*

1  
2  
3 Samples collected before 35 weeks' GA were selected to build the model to predict PTB.  
4  
5 Mann–Whitney U test was used to select the initial candidate metabolites that were then  
6  
7 mapped to pathways. The value of each pathway was calculated as the weighted sum of  
8  
9 the normalized concentrations of metabolites on the pathway divided by the number of  
10  
11 metabolites. The weight of each metabolite was the absolute value of the ratio of median  
12  
13 of full-term samples to PTB samples. Like the GA estimation, pathways could have two  
14  
15 values that depended on the ratio of median greater or less than 1. An XGBoost model  
16  
17 was developed utilizing samples from the SU cohort and validated with the UAB cohort.  
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

## Text A.2 Metabolite model vs. IBP4/SHBG in predicting PTB

We conducted ELISA tests to evaluate the IBP4/SHBG signature, a predictor that was validated in a prospective study as a predictor of spontaneous PTB. Commercial kits Human IGFBP4 ELISA Kit (Abcam, Burlingame, CA, USA) and Human SHBG Quantikine ELISA Kit (R&D System Inc.) were used. AUC of the predictor was calculated in different GA intervals and with different maternal BMI values, and was compared to the performance of the metabolic model.

With a BMI of  $>22$  and  $\leq 37$  kg/m<sup>2</sup>, the AUC values of the IBP4/SHBG predictor peaked at 15–20 weeks' GA (SU: 0.833; UAB: 1), and dropped rapidly after 20 weeks (Figure A below). The AUC values were lower with extreme BMI (0.7 at BMI  $\leq 20$  kg/m<sup>2</sup> and 0.63 at BMI  $>27$  kg/m<sup>2</sup>; see Figure B below). These findings are consistent with the previous validation study. Compared with the IBP4/SHBG predictor, the metabolic model has a more stable AUC performance over the gestation and different BMI values in SU ( $P = 0.03$ ). In UAB at  $>18$  weeks' GA, the AUC of IBP4/SHBG dropped from 0.6 to 0.3, while the AUC of the metabolic model was above 0.8.

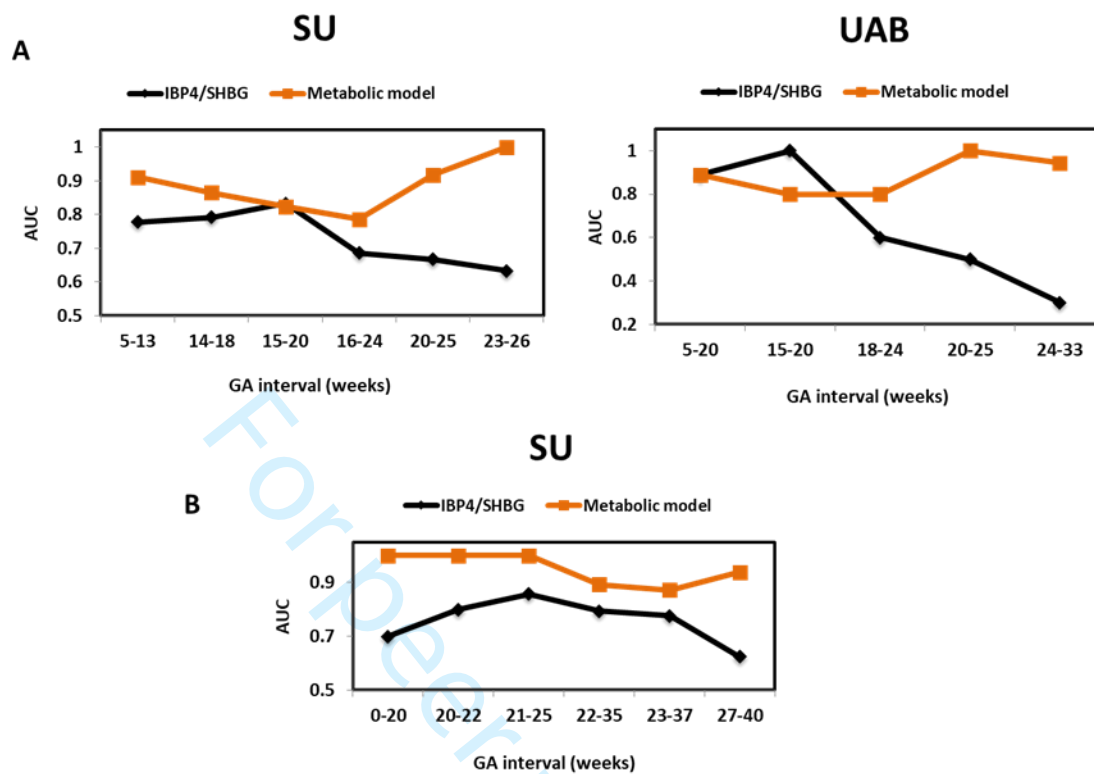


Figure. The performance of the IBP4/SHBG predictor and the metabolic model. The results are stratified by the GA intervals with a BMI at 22–37 kg/m<sup>2</sup> (A), and by BMI values with a GA interval of 5–20 weeks (B).

## STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	Page No
<b>Title and abstract</b>	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	1
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	2-3
<b>Introduction</b>			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	5
Objectives	3	State specific objectives, including any prespecified hypotheses	6
<b>Methods</b>			
Study design	4	Present key elements of study design early in the paper	6
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6,7
Participants	6	(a) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up <i>Case-control study</i> —Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls <i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants	6,7
		(b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed <i>Case-control study</i> —For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6,7,8
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	7
Bias	9	Describe any efforts to address potential sources of bias	8
Study size	10	Explain how the study size was arrived at	
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	7,8
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	8
		(b) Describe any methods used to examine subgroups and interactions	
		(c) Explain how missing data were addressed	
		(d) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed <i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy	
		(e) Describe any sensitivity analyses	

Continued on next page



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

<b>Results</b>			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	9
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	9
		(b) Indicate number of participants with missing data for each variable of interest	
		(c) <i>Cohort study</i> —Summarise follow-up time (eg, average and total amount)	9
Outcome data	15*	<i>Cohort study</i> —Report numbers of outcome events or summary measures over time	10
		<i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure	
		<i>Cross-sectional study</i> —Report numbers of outcome events or summary measures	
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	11-12
		(b) Report category boundaries when continuous variables were categorized	11-12
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	12
<b>Discussion</b>			
Key results	18	Summarise key results with reference to study objectives	13
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	16
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	14-15
Generalisability	21	Discuss the generalisability (external validity) of the study results	16
<b>Other information</b>			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	18

\*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

**Note:** An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at [www.strobe-statement.org](http://www.strobe-statement.org).

# BMJ Open

## A metabolic clock as noninvasive blood tests of preterm birth and for gestational age assessment: a two-center retrospective study in the US

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2020-040647.R1
Article Type:	Original research
Date Submitted by the Author:	19-Oct-2020
Complete List of Authors:	Sylvester, Karl; Stanford University Hao, Shiyong; Stanford University You, Jin; Stanford University Zheng, Le; Stanford University Tian, Lu; Stanford University Yao, Xiaoming; West China Hospital of Medicine Mo, Lihong; University of California San Francisco Fresno Center for Medical Education and Research Ladella, Subhashini; University of California San Francisco Fresno Center for Medical Education and Research Wong, Ronald; Stanford University Shaw, Gary M.; Stanford University Stevenson, David; Stanford University Cohen, Harvet; Stanford University Whitin, John; Stanford University McElhinney, Doff ; Stanford University Ling, Xuefeng; Stanford University,
<b>Primary Subject Heading</b>:	Obstetrics and gynaecology
Secondary Subject Heading:	Paediatrics
Keywords:	Health informatics < BIOTECHNOLOGY & BIOINFORMATICS, Risk management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, OBSTETRICS

SCHOLARONE™  
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

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

1 **A metabolic clock as noninvasive blood tests of preterm birth and for**  
2 **gestational age assessment: a two-center retrospective study in the US**

3 Karl G. SYLVESTER, MD<sup>1\*†</sup>, Shiyong HAO, PhD<sup>2,3\*</sup>, Jin YOU, PhD<sup>1\*</sup>, Le ZHENG,  
4 PhD<sup>2,3</sup>, Lu TIAN<sup>4</sup>, Xiaoming YAO, MD, PhD<sup>5</sup>, Lihong MO, MD<sup>6</sup>, Subhashini  
5 LADELLA, MD<sup>6</sup>, Ronald J. WONG, BA<sup>7</sup>, Gary M. SHAW, DrPh<sup>7</sup>, David K.  
6 STEVENSON, MD<sup>7</sup>, Harvey J. COHEN, MD, PhD<sup>7</sup>, John C. WHITIN, PhD<sup>7</sup>, Doff B.  
7 MCELHINNEY, MD<sup>2,3</sup>, Xuefeng B. LING, PhD<sup>1,3†</sup>

8  
9 <sup>1</sup>Department of Surgery, Stanford University School of Medicine, Stanford, CA, USA.

10 <sup>2</sup>Department of Cardiothoracic Surgery, Stanford University School of Medicine,  
11 Stanford, CA, USA.

12 <sup>3</sup>Clinical and Translational Research Program, Betty Irene Moore Children's Heart Center,  
13 Lucile Packard Children's Hospital, Palo Alto, CA, USA.

14 <sup>4</sup>Department of Health Research and Policy, Stanford University, Stanford, CA, USA.

15 <sup>5</sup>Translational Medicine Laboratory, West China Hospital, Chengdu, China.

16 <sup>6</sup>Department of Obstetrics and Gynecology, University of California San Francisco-  
17 Fresno, Fresno, CA, USA.

18 <sup>7</sup>Department of Pediatrics, Stanford University School of Medicine, Stanford, CA, USA.

19  
20 †Correspondence: Xuefeng B. Ling, Department of Surgery, Stanford University,  
21 bxling@stanford.edu; Karl G. Sylvester, Department of Surgery, Stanford University,  
22 karls@stanford.edu

23 \*Co-first authors

24  
25 **Word count:** 293 for Abstract; 3508 for Main text.

1  
2  
3 26 **ABSTRACT**  
4  
5

6 27 **Objectives** The aim of this study was to develop a single blood test that could determine  
7  
8 28 gestational age and estimate the risk of preterm birth by measuring serum metabolites.  
9  
10 29 We hypothesized that serial metabolic modeling of serum analytes throughout pregnancy  
11  
12 30 could be used to describe fetal gestational age and project preterm birth with a high  
13  
14 31 degree of precision.  
15  
16

17  
18 32 **Study design** A retrospective cohort study  
19

20  
21 33 **Setting** Two medical centers from US  
22

23  
24 34 **Participants** Thirty-six patients (20 full-term, 16 preterm) enrolled at Stanford  
25  
26 35 University were used to develop gestational age and preterm birth risk algorithms, 22  
27  
28 36 patients (9 full-term, 13 preterm) enrolled at the University of Alabama were used to  
29  
30 37 validate the algorithms.  
31  
32

33 38 **Outcome measures** Maternal blood was collected serially throughout pregnancy.  
34

35  
36 39 Metabolic datasets were generated using mass spectrometry.  
37

38 40 **Results** A model to determine gestational age was developed ( $R^2 = 0.98$ ) and validated  
39  
40 41 ( $R^2 = 0.81$ ). 66.7% of the estimates fell within  $\pm 1$  week of ultrasound results during  
42  
43 42 model validation. Significant disruptions from full-term pregnancy metabolic patterns  
44  
45 43 were observed in preterm pregnancies ( $R^2 = -0.68$ ). A separate algorithm to predict  
46  
47 44 preterm birth was developed utilizing a set of 10 metabolic pathways that resulted in an  
48  
49 45 area under the curve of 0.96 and 0.92, a sensitivity of 0.88 and 0.86, and a specificity of  
50  
51 46 0.96 and 0.92 during development and validation testing, respectively.  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 47 **Conclusions** In this study metabolic profiling was used to develop and test a model for  
4  
5 48 determining gestational age during full-term pregnancy progression, and to determine  
6  
7 49 risk of preterm birth. With additional patient validation studies, these algorithms may be  
8  
9 50 used to identify at-risk pregnancies prompting alterations in clinical care, and to gain  
10  
11 51 biologic insights into the pathophysiology of preterm birth. Metabolic pathway-based  
12  
13 52 pregnancy modeling is a novel modality for investigation and clinical application  
14  
15 53 development.  
16  
17  
18  
19

20 54 **Keywords:** Metabolic, gestational age, preterm birth, pathway  
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

1  
2  
3 **56 Strengths and limitations of this study**  
4  
5

- 6 57 • The insensitivity of the prediction model to gestational age (GA) window of sample  
7  
8 58 collection increases its flexibility and opportunity for potential clinical use.  
9  
10 59 • This study is among the first to propose a pathway-based computational methodology  
11  
12 to estimate GA and predict preterm birth.  
13  
14 60  
15 61 • The overall cohort size is modest, and the distribution of sampling time are different  
16  
17 between patients and cohorts.  
18  
19 62  
20 63 • It is a retrospective study; a larger prospective cohort study is necessary before  
21  
22 applying the estimates and prediction to a broader population for clinical utility.  
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

## 65 INTRODUCTION

66 Gestational age (GA) dating is a core element of standard prenatal care<sup>1-4</sup>. Prenatal  
67 ultrasound (US) is an established modality for estimating GA, monitoring fetal growth,  
68 and screening for fetal anomalies<sup>5</sup>. According to the policy statement of the Committee  
69 on Obstetric Practice, the American Institute of Ultrasound in Medicine, and the Society  
70 for Maternal-Fetal Medicine, a pregnancy is considered optimally dated through a  
71 combination of last menstrual period (LMP) and an accurate US obtained prior to 22 0/7  
72 weeks<sup>6</sup>. Accordingly, LMP is dependent on maternal recall and many pregnancies do not  
73 present for a first prenatal US evaluation until the second or third trimester. Thus, there is  
74 a need for a molecular method that would complement the potential shortcomings of  
75 LMP recall and US dating outside the first trimester. Moreover, it is possible that  
76 molecular pregnancy dating will provide greater resolution to pregnancy risk than current  
77 information based on calendar dating (LMP) and anthropometrics (US). Although  
78 experience is accumulating with the use of second and third trimester US for an  
79 estimation of risk of preterm birth (PTB)<sup>7-9</sup>, to date these measures have not been widely  
80 adopted, are subject to user experience and have reported variable performance  
81 characteristics. The availability and expertise of US in disadvantaged areas is limited<sup>10</sup>.  
82 Therefore, there is a need to develop an alternative measure of fetal progression to  
83 estimate GA and pregnancy risk in a variety of settings and especially when US and LMP  
84 dates are unavailable or unreliable.

85 Compared with imaging methodologies, blood-based molecular testing may provide a  
86 more reproducible and precise modality in clinical applications for the frequent  
87 monitoring of health status and detection of early signs of disease. Genomic, gene



1  
2  
3 88 expression, protein, and metabolite profiles measured in human blood have been  
4  
5 89 increasingly utilized for the determination of disease risk and to gain disease specific  
6  
7 90 pathophysiology insight. Attempts at estimating GA using molecular adaptations have  
8  
9  
10 91 included modeling of RNA, protein, or immune cell changes, and most recently  
11  
12 92 metabolites in maternal or newborn blood <sup>11-17</sup>. Similarly, risk prediction of PTB in  
13  
14 93 clinical settings is currently primarily based on maternal history. Biomarkers have been  
15  
16 94 suggested from genetic and proteomic analyses, but less effort has been focused on  
17  
18  
19 95 understanding maternal metabolic signatures of pregnancy <sup>18-24</sup>.

20  
21  
22 96 In this study, we hypothesized that longitudinal metabolic profiling of pregnancy reflects  
23  
24 97 the temporal progression of fetal development with a high degree of precision. Moreover,  
25  
26 98 we posited that if a normal pregnancy progression profile could be defined in metabolic  
27  
28 99 terms, then aberrations from the normal profile may identify a pregnancy at risk for PTB.  
29  
30  
31 100 Our findings suggest that composite metabolic panel modeling may serve as a  
32  
33 101 reproducible and precision approach to GA dating of pregnancy and prediction of PTB.  
34  
35

## 36 102 **MATERIALS AND METHODS**

### 37 38 39 103 **Definition**

40  
41  
42 104 In this study, a full-term pregnancy was defined as a pregnancy ending with a delivery at  
43  
44 105  $\geq 37$  weeks. PTB was defined by delivery at  $< 35$  weeks GA in order to make a complete  
45  
46 106 separation from the full-term subjects.  
47  
48

### 49 107 **Study design**

50  
51  
52 108 The study was conducted in two phases: (1) modeling to devise a metabolite-based  
53  
54 109 estimation of GA during full-term pregnancies; and (2) modeling to devise a metabolic  
55  
56  
57  
58  
59  
60

1  
2  
3 110 panel predictive of PTB (Fig. 1). In this study, the ‘gold’ standard of GA was US  
4  
5 111 measurement based on the crown-rump length at the first trimester <sup>25</sup>. Serum samples  
6  
7 112 were collected in the 1<sup>st</sup>, 2<sup>nd</sup>, or 3<sup>rd</sup> trimester during pregnancy for each individual  
8  
9 113 woman. Each participant had 1 to 4 time-points collected prior to delivery. Samples were  
10  
11 114 provided by Stanford Hospital and Clinics (SU) and the University of Alabama (UAB).  
12  
13 115 Metabolic concentrations in each sample were measured by targeted and untargeted mass  
14  
15 116 spectrometry (MS) analysis. Models that estimated GA or predicted PTB were developed  
16  
17 117 using the SU cohort and validated using the UAB cohort. The study was approved by the  
18  
19 118 Institutional Review Board of both sites (Protocol #21956). All samples were collected  
20  
21 119 after informed consent was obtained. All statistical analyses were done in R software.  
22  
23  
24  
25

### 26 120 **Targeted and global MS analysis**

27  
28  
29 121 Samples of full-term and preterm patients as well as quality control (QC) samples were  
30  
31 122 injected into the MS. Targeted MS analysis was done through flow injection methods by  
32  
33 123 using Ultimate 3000 Ultra-High-Performance Liquid Chromatography (UHPLC) system  
34  
35 124 and Quantiva Triple Quadrupole Mass Spectrometer. Global (i.e. untargeted) MS analysis  
36  
37 125 was done by using a Vanquish UHPLC system coupled to a Q Exactive plus mass  
38  
39 126 spectrometer and Q Exactive HF hybrid quadrupole-Orbitrap mass spectrometer.  
40  
41  
42  
43

### 44 127 **Data preprocessing and metabolic identification**

45  
46 128 A data pre-processing procedure was conducted to convert the raw data generated by MS  
47  
48 129 analysis into a matrix of relative concentrations of metabolites versus samples <sup>26</sup>. This  
49  
50 130 procedure was done by R package. Metabolic values in each sample were then  
51  
52 131 normalized by the median values measured with QC samples to reduce the batch effects.  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 132 Compounds detected by untargeted analyses were matched to metabolites in the Human  
4  
5 133 Metabolome Database by putative identification <sup>27</sup>. Accurate mass was used for the  
6  
7  
8 134 mapping. Metabolites were mapped to pathways using Kyoto Encyclopedia of Genes and  
9  
10 135 Genomes (KEGG) and Human Metabolome Database (HMDB). Only endogenous  
11  
12 136 pathways were considered.

### 137 **Metabolic compound selection, pathway computation, and model development**

138 Metabolites measured by targeted and untargeted MS were aggregated and filtered. The  
139 remaining metabolites were mapped to pathways. The value of each pathway was  
140 calculated as the weighted sum of the normalized concentrations of metabolites on the  
141 pathway divided by the number of metabolites. An XGBoost model was developed with  
142 the pathway values of samples from full-term patients to estimate the GA. R-squared ( $R^2$ ;  
143 goodness-of-fit of the model), root-mean-square error (RMSE), and error distribution  
144 were calculated to evaluate the model performance. A second XGBoost model was  
145 developed to predict PTB. To evaluate the model performance, Mann–Whitney U tests  
146 were used to compare the distribution of final predictive estimates, i.e., XGBoost model  
147 values, on full-term and PTB samples. Additional details of model development were  
148 described in Text A.1. ELISA tests were conducted on the SU and UAB cohorts to  
149 evaluate the insulin-like growth factor-binding protein 4 (IBP4)/sex hormone-binding  
150 globulin (SHBG) signature, a predictor that was validated in a prospective study as a  
151 predictor of spontaneous PTB <sup>19</sup>. Serum concentrations were measured using commercial  
152 kits Human IGFBP4 ELISA Kit (Abcam, Burlingame, CA, USA) and Human SHBG  
153 Quantikine ELISA Kit (R&D System Inc.). Results were compared with our metabolic  
154 model.

## 155 Patient and Public Involvement statement

156 This retrospective research was done without patient involvement. Patients were not  
 157 invited to comment on the study design and were not consulted to develop patient  
 158 relevant outcomes or interpret the results. Patients were not invited to contribute to the  
 159 writing or editing of this document for readability or accuracy.

## 160 RESULTS

### 161 Samples

162 As shown in Fig. 2, the SU cohort had 20 full-term pregnancies with 57 blood samples  
 163 (17, 32, and 8 collected in the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> trimesters, respectively) and 16 preterm  
 164 pregnancies with 32 blood samples (9, 19, and 4 collected in the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup>  
 165 trimesters, respectively). The UAB cohort had 9 full-term pregnancies with 13 blood  
 166 samples (8 and 5 in the 2<sup>nd</sup>, and 3<sup>rd</sup> trimesters, respectively) and 13 preterm pregnancies  
 167 with 22 blood samples (4 and 18 in the 1<sup>st</sup> and 2<sup>nd</sup> trimesters, respectively). In the SU  
 168 cohort, 2 (12.5%) were extremely preterm (< 28 weeks), and 5 (31.3%) were very  
 169 preterm (28–31 weeks). In the UAB cohort, 6 (46.2%) were extremely preterm, and 3  
 170 (23.1%) were very preterm. Our SU and UAB cohorts were assembled: no complications  
 171 of pregnancy were included; all deliveries were singleton; and all PTB were spontaneous.  
 172 Demographics of the two cohorts are shown in Table 1.

173 **Table 1.** Maternal characteristics in SU and UAB cohorts

Characteristic	SU			UAB			SU vs.
	Full-term	Preterm	<i>P</i>	Full-term (n =	Preterm	<i>P</i>	UAB
							<i>P</i>

	(n = 20)	(n = 16)		9)	(n = 13)		
Race, n (%)			<b>&lt;0.001</b>			0.5	<b>&lt;0.001</b>
Asian	0	1 (6.3)		0	0		
White	20 (100)	5 (31.3)		0	2 (15.4)		
Black	0	1 (6.3)		9 (100)	10 (76.9)		
American Indian	0	2 (12.5)		0	0		
Pacific Islander	0	1 (6.3)		0	0		
Other/unknown	0	6 (37.5)		0	1 (7.7)		
Hispanic, n (%)	0	8 (50)	<b>&lt;0.001</b>	0	1 (7.7)	0.9	0.1
Maternal Age, year, mean (SD)	31.9 (4.8)	29.8 (7.5)	0.3	25.6 (5.0)	27.5 (4.5)	0.4	<b>0.008</b>
Gestational age at delivery, weeks, median (IQR)	39.5 (39,41)	32 (30,33)	<b>&lt;0.001</b>	38 (37,39)	28 (26,32)	<b>&lt;0.001</b>	<b>0.01</b>
Having previous pregnancy, n (%)	9 (45)	6 (37.5)	0.7	9 (100)	13 (100)	0.4	<b>&lt;0.001</b>
BMI, kg/m <sup>2</sup> , median (IQR)	22.3 (20.2,24.7)	27.6 (23.4,33.9)	<b>0.003</b>	30.4 (22.3,33.1)	26.5 (22.6,36.5)	0.8	0.06
History of PTB, n (%)	3 (15)	8 (50)	<b>0.03</b>	7 (77.8)	13 (100)	0.2	<b>&lt;0.001</b>

175

176

177 **LC-MS/MS metabolomics**

178 The study targeted 315 metabolites by LC-MS/MS, including 13 categories: acyl-

179 carnitine (11, 3.5%), amino acid (9, 2.9%), fatty acid (6, 1.9%), ceramide (12, 3.8%),

180 ceramide 1-phosphate (8, 2.5%), galactosylceramide (5, 1.6%), phosphatidyl acid (15,

1  
2  
3 181 4.8%), phosphatidylethanolamine (52, 16.5%), phosphatidylglycerol (5, 1.6%),  
4  
5 182 phosphatidylinositol (11, 3.5%), phosphatidylcholine (130, 41.3%), cholesteryl ester (16,  
6  
7 183 5.1%), and sphingomyelin (35, 11.1%). The study also identified 1627 positively-and 295  
8  
9 184 negatively-charged compounds through untargeted analyses. Together these formed the  
10  
11 185 initial set of 2237 compounds.  
12  
13  
14

### 15 186 **Feature selection of GA estimation modeling**

16  
17  
18 187 Of the 2237 compounds, 118 had an absolute Pearson correlation coefficient of  $> 0.35$   
19  
20 188 with GA. The cutoff of  $\pm 0.35$  was selected based on the false discovery rate (FDR)  
21  
22 189 values of the mapped pathways  $< 1\%$  (Fig. A.1). The 118 compounds were mapped to 89  
23  
24 190 pathways, 33 of which were selected by the XGBoost model. The normalized value of  
25  
26 191 each pathway varied over the course of gestation (Fig. A.2). Univariate analysis of the 33  
27  
28 192 pathways is shown in Fig. A.3, and the top 10 pathways in the model is depicted in Fig. 3.  
29  
30 193 The top 10 pathways included those associated in the metabolisms of:  
31  
32 194 glycerophospholipid, arginine and proline, thiamine, purine, butanoate, galactose, sulfur,  
33  
34 195 phenylalanine, and C5-branched dibasic acid.  
35  
36  
37  
38

### 39 196 **Performance of GA estimation**

40  
41  
42 197 The performance of GA estimates on full-term samples was similar in the development  
43  
44 198 phase (SU cohort,  $R^2 = 0.98$ , RMSE = 1.09) and the validation phase (UAB cohort,  $R^2 =$   
45  
46 199  $0.81$ , RMSE = 2.36) (Fig. 4). In our validation testing, 66.7% of the estimates were  
47  
48 200 within  $\pm 1$  week of the US results (Fig. A.4).  
49  
50

51  
52 201 Intriguingly, model performance significantly deteriorated when applied to samples from  
53  
54 202 PTB pregnancies ( $R^2 = -0.68$  and RMSE = 6.6 in validation; see Fig. 4). It suggested that  
55  
56  
57  
58  
59  
60

1  
2  
3 203 the relationships between metabolic parameters and full-term pregnancies were not  
4  
5 204 maintained in PTB pregnancies. Furthermore, such disruptions were notable as early as  
6  
7 205 10 weeks' GA (Fig. 4) or early to mid-gestation. These findings prompted the  
8  
9  
10 206 development of a metabolic-based model of PTB estimation.

### 11 12 13 207 **Performance of PTB prediction**

14  
15 208 Samples collected before 35 weeks' GA were used to develop a model that differentiated  
16  
17 209 PTB pregnancies from those full-term. As before, the model was developed with the SU  
18  
19 210 cohort that had 20 full-term (54 samples) and 16 preterm (32 samples) pregnancies, and  
20  
21 211 was validated with the UAB cohort that had 9 full-term (13 samples) and 13 preterm (22  
22  
23 212 samples) pregnancies. In total, 148 metabolic compounds (with Mann-Whitney U test  $P <$   
24  
25 213 0.05) were mapped to 66 pathways (FDR  $< 1.5\%$ ; see Fig. A.5). Further model  
26  
27 214 development selected 10 pathways as strong predictors covering the metabolisms of  
28  
29 215 glycerophospholipid, sphingolipid, taurine and hypotaurine, arachidonic acid, secondary  
30  
31 216 bile acid biosynthesis, glycerolipid, cysteine and methionine, tryptophan, and arginine  
32  
33 217 and proline (Fig. 5).

34  
35  
36  
37  
38  
39 218 The level of prediction accuracy was maintained in the validation cohort ( $P = 5 \times 10^{-5}$ , area  
40  
41 219 under the curve [AUC] = 0.92; see Fig. 6). The prevalence-corrected positive predictive  
42  
43 220 values (PPVs) across model values (*i.e.* scores) were plotted based on the PTB  
44  
45 221 prevalence in Alabama in 2018 (12.5%; see Fig. A.6). A threshold value of 0.52 was  
46  
47 222 selected as a high-risk threshold for PTB, which was associated with a PPV of 0.70, a  
48  
49 223 relative risk (RR) of 5.6 compared to the United States population baseline (=   
50  
51 224 0.70/12.5%), a sensitivity of 0.86 (19 of 22), and a specificity of 0.92 (12 of 13; Fig. 7).  
52  
53  
54  
55 225 The sensitivities and specificities with cutoff values are shown in Table A.1.

1  
2  
3 226 In the validation cohort, 12 of 13 full-term samples and 19 of 22 preterm samples were  
4  
5 227 classified correctly. The misclassified full-term sample was from a mother that delivered  
6  
7 228 at 37 weeks' GA. The 19 correctly classified PTB samples were from 13 PTB  
8  
9 229 pregnancies. Of the 13 pregnancies, 9 were identified as high risk at or earlier than 16  
10  
11 230 weeks' GA. The median gap between the time of identification and the delivery was 11  
12  
13 231 weeks' GA (IQR: 8, 15.5).

14  
15  
16  
17 232 To determine the performance of our metabolic model against existing models, a  
18  
19 233 comparison between the metabolic PTB risk model and the commercially available  
20  
21 234 IBP4/SHBG PTB test was performed and summarized in Text A.2 and Fig. A.7.

### 22 23 24 25 235 **Metabolite-based model and pathway-based model: a comparison**

26  
27  
28 236 To determine the effectiveness of model performance based upon robustness of biologic  
29  
30 237 features, we compared model performance using pathway or individual metabolite as  
31  
32 238 selected features in estimating GA and predicting PTB. The performance of the pathway-  
33  
34 239 based models were significantly better than the metabolite-based models, with a lower  
35  
36 240 RMSE (Student's t-test  $P = 4 \times 10^{-3}$ ; Fig. A.8) and a larger AUC (DeLong test  $P = 0.03$ ;  
37  
38 241 Fig. A.9).

## 39 40 41 42 242 **DISCUSSION**

### 43 44 45 243 **Principal Findings**

46  
47  
48 244 In this study, we report a panel of metabolic pathways measured in maternal serum that  
49  
50 245 provides an estimation of GA over the course of a full-term pregnancy. A second and  
51  
52 246 distinct set of metabolic pathways was also identified in maternal serum that could  
53  
54 247 distinguish pregnancies ending with PTB (< 35 weeks) from full-term ( $\geq 37$  weeks) with



1  
2  
3 248 a high degree of precision. The models were developed and validated using two  
4  
5 249 independent cohorts from two different institutions in order to test the robustness of the  
6  
7 250 biologic features driving the classifications. Intriguingly, PTB pregnancies do not  
8  
9 251 demonstrate the same temporal relationship as term pregnancies upon metabolic  
10  
11 252 modeling across gestation (Fig. 4). Indeed, PTB pregnancies demonstrate a marked  
12  
13 253 departure from the term metabolic profile (Fig. 4) that is not only dramatic ( $R^2= 0.98$   
14  
15 254 train and 0.81 test for term model; compared to  $R^2= 0.50$  train and -0.68 test for PTB  
16  
17 255 pregnancy in term model), but is also recognizable as early as 10 weeks' GA as  
18  
19 256 determined by the current standard of US dating. Recognizing the metabolic pathway  
20  
21 257 aberration of PTB pregnancies, a second model was developed using metabolic pathway  
22  
23 258 analyses to quantify the risk of PTB prior to 35 weeks' GA. Once again, metabolic  
24  
25 259 profiling proved to be robust in identifying PTB pregnancies with a high degree of  
26  
27 260 sensitivity (AUC 0.96 training; AUC 0.92 testing) and precision (training PPV 0.93  
28  
29 261 (0.78-0.99); testing PPV 0.95 (0.75-1). Taken together, this study demonstrated a  
30  
31 262 powerful new, reproducible methodology for monitoring pregnancy progression and  
32  
33 263 identifying abnormal pregnancies.  
34  
35  
36  
37  
38  
39

#### 40 264 **Clinical and Research Implications**

41  
42  
43 265 The potential clinical utility of developing a test for pregnancy monitoring is appealing.  
44  
45 266 There is a need to develop a more robust method than LMP and an alternative to first  
46  
47 267 trimester US that captures pregnancy progression, a complex relationship of fetal and  
48  
49 268 placental growth, development, and function. To support these processes, there is a need  
50  
51 269 for energy transfer between mother and fetus throughout gestation. We therefore  
52  
53 270 reasoned that metabolic phenotyping would be ideally suited to capture this relationship.  
54  
55  
56  
57  
58  
59

1  
2  
3 271 Despite a modest cohort size, the results of metabolic modeling demonstrate a high  
4  
5 272 degree of concordance with clinical standard US dating performed by experts as reflected  
6  
7 273 by 66.7% of model estimates falling within  $\pm 1$  week of US results (Fig. A.4). Moreover,  
8  
9 274 unlike the deterioration experienced with US dating of pregnancy, metabolic modeling  
10  
11 275 was shown to achieve near equivalent performance in the 1<sup>st</sup>, 2<sup>nd</sup>, and 3<sup>rd</sup> trimesters,  
12  
13 276 indicating the potential for broad clinical applicability that might achieve independence  
14  
15 277 of reliance on accuracy of LMP or concordance among modality testing. The result of  
16  
17 278 PTB prediction is equally robust demonstrating a high degree of precision. Beyond  
18  
19 279 relying on clinical histories or self-reported symptoms, the model proposed here provides  
20  
21 280 a molecular classification that may be more accurate than current methods and further  
22  
23 281 reflect a comprehensive measure of aberrant pregnancy based on metabolic changes. In  
24  
25 282 practice, clinicians could use the PTB prediction model to differentiate high- from low-  
26  
27 283 risk patients. Low risk patients would then be subject to GA estimation panel testing, all  
28  
29 284 from the same blood draw.

30  
31  
32  
33  
34  
35 285 A distinct advantage of the PTB risk prediction developed in this study is that it has a  
36  
37 286 wide window of sampling. Samples were collected broadly before 35 weeks' GA, which  
38  
39 287 is wider than the window of other well-established biomarkers such as fetal fibronectin  
40  
41 288 (between 24 and 34 weeks' GA)<sup>20</sup>, IBP4/SHBG (19 to 21 weeks)<sup>19</sup>, and inter-alpha-  
42  
43 289 trypsin inhibitor heavy chain 4 protein (24 and 28 weeks)<sup>18</sup>. Relatively stable AUC  
44  
45 290 levels were maintained throughout the diagnostic window (Text A.2). The insensitivity of  
46  
47 291 the prediction model to GA at testing increases its flexibility and opportunity for potential  
48  
49 292 clinical use. An additional advantage of the model herein is the ability for early  
50  
51 293 identification of high-risk women. Although there is no standardized guideline for early-  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 294 gestation management of patients at risk of PTB delivery, metabolic modeling for PTB  
4  
5 295 risk may provide a not previously possible opportunity for early gestation risk mitigation.  
6  
7  
8 296 Clinical trials have suggested that hormone treatment and maternal physical activity  
9  
10 297 modifications applied between 16 to 37 weeks' GA reduced the PTB rate of women who  
11  
12 298 were deemed at high risk due to a history of prior PTB delivery<sup>28 29</sup>. In many cases PTB  
13  
14 299 can not be prevented, however any opportunity is deemed highly desirable for even a  
15  
16 300 modest delay (1–2 weeks) in PTB or an enhanced ability to more accurately triage for  
17  
18 301 delivery to centers with the capability to manage profoundly premature neonates<sup>30-32</sup>.  
19  
20  
21  
22 302 This study is among the first to propose a pathway-based computational methodology to  
23  
24 303 estimate GA and predict PTB. Metabolic pathways are linked to chemical functions, and  
25  
26 304 the alteration or disruption of specific functions participate in disease phenotypes,  
27  
28 305 facilitating the use of pathways to function as higher-level biomarkers of diseases<sup>33</sup>. The  
29  
30 306 role of metabolic pathways in disease diagnosis has been explored in several preliminary  
31  
32 307 clinical studies<sup>34 35</sup>. Pathway performance in differentiating patients with disease from  
33  
34 308 healthy controls has been found to be effective compared to using individual metabolites  
35  
36 309<sup>35</sup>. Similarly, we found the pathway-based models had less variability and higher  
37  
38 310 sensitivity than metabolite-based models that were developed using the same population.  
39  
40  
41  
42 311 One plausible explanation for this observation may be attributed to the calculation of  
43  
44 312 pathway values, which represents the sum of individual metabolites and thus may  
45  
46 313 amplify association to outcome relationships. This hypothesis is supported by the FDR  
47  
48 314 comparison (Fig. A.8 and A.9): pathway-based analysis had lower FDR values than  
49  
50 315 metabolite models. This study adds to the exploration of the feasibility of using pathways  
51  
52  
53 316 for health monitoring and prediction.  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 317 In this study glycerophospholipid metabolism was identified as the most significant  
4  
5 318 contributing pathway for both gestational age estimation and preterm birth prediction.  
6  
7 319 Glycerophospholipids consist of fatty acid chains and have been previously cited as  
8  
9 320 strong correlates to birth weight, pregnancy duration, and risk of preterm birth<sup>36</sup>. These  
10  
11 321 same authors also found different polyunsaturated fatty acid components of  
12  
13 322 glycerophospholipid had differential effects on fetal growth. Gao et al has reported a  
14  
15 323 potential association between glycerophospholipid and labor timing in rodent models<sup>37</sup>  
16  
17 324<sup>38</sup>. The current study extends those prior observations through a quantitative assessment  
18  
19 325 of the relationship between glycerophospholipid metabolism, gestational age and the risk  
20  
21 326 of preterm birth. The leading effect of glycerophospholipid pathway metabolism in the  
22  
23 327 current study was positive in both the assessment of gestational age and risk of preterm  
24  
25 328 birth. These findings add further insight into the role of glycerophospholipid metabolism  
26  
27 329 in human pregnancy. Other contributing pathways for preterm birth prediction such as  
28  
29 330 sphingolipid metabolism, arachidonic acid metabolism, and arginine and proline  
30  
31 331 metabolism were also found associated to preterm. Alterations in plasma sphingolipids  
32  
33 332 were found in women who had spontaneous PTB<sup>39</sup>. Increase of arachidonic acid  
34  
35 333 metabolism might correlate to bacteria activities that led to preterm labor<sup>40</sup>. Plasma level  
36  
37 334 of arginine and citrulline was significantly lowered in preterm babies<sup>41</sup>.  
38  
39  
40  
41  
42  
43  
44  
45 335 Taken together, the analysis of the leading pathways found to significantly contribute to  
46  
47 336 the metabolic pregnancy modeling herein provide ample insights to deepen our  
48  
49 337 understanding of pregnancy progression and may facilitate the identification and  
50  
51 338 interpretation of potential therapeutic targets. Further, we speculate that the platform and  
52  
53 339 approaches outlined herein may be extended to the interrogation of additional conditions  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 340 of pregnancy including abnormalities of placentation, gestational diabetes and fetal  
4  
5 341 growth disturbances among others.  
6  
7

### 8 342 **Limitations**

9  
10  
11 343 This study has several limitations. First, the overall cohort size was modest, and  
12  
13 344 pregnancies with delivery at 35 or 36 weeks were not included in the study. Second,  
14  
15 345 blood samples were collected in a non-uniform manner with respect to GA timing and  
16  
17 346 time of day. The time between two adjacent samples corresponding to the same patient  
18  
19 347 varied. Third, the distribution of samples throughout pregnancy were different between  
20  
21 348 patients and cohorts. In the SU cohort, none of the full-term patients had samples  
22  
23 349 collected between 30 and 37 weeks. In the UAB cohort, none of the full-term patients had  
24  
25 350 sampling in the 1<sup>st</sup> trimester, and none of the PTB patients had sampling in the 3<sup>rd</sup>  
26  
27 351 trimester. Fourth, for methodologic reasons, not all serum analytes could be identified  
28  
29 352 and mapped to known metabolites. Fifth, baseline characteristics of patients were not  
30  
31 353 included in the analysis. Sixth, the study was retrospective, and the participants were  
32  
33 354 solely from California and Alabama. A larger prospective cohort study with a reasonable  
34  
35 355 ratio of full-term to preterm is necessary before applying the estimates and prediction to a  
36  
37 356 broader population for clinical utility.  
38  
39  
40  
41  
42  
43

### 44 357 **CONCLUSION**

45  
46 358 The present study demonstrates that maternal serum based metabolic profiling is a highly  
47  
48 359 sensitive and accurate method for determining GA and prediction of PTB. The pathway-  
49  
50 360 based analysis supports the hypothesis of the orderly metabolic progression of pregnancy  
51  
52 361 that can be reproducibly captured using metabolic profiling. The robustness of the  
53  
54 362 modeling reinforces the potential appeal for further clinical development and as a  
55  
56  
57  
58  
59  
60

1  
2  
3 363 platform to investigate the pathophysiology associated with aberrant fetal development  
4  
5 364 and pregnancy progression. This study is the first to report a single blood test for  
6  
7  
8 365 metabolic pathway-based determination of GA dating, and early detection of PTB risk.  
9

10 366  
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

For peer review only

1  
2  
3 367 **Acknowledgments:** The authors thank colleagues at the Stanford University Pediatric  
4  
5 368 Proteomics group and the March of Dimes Prematurity Research Center at Stanford  
6  
7  
8 369 University for critical discussions.

9  
10 370 **Funding:** This work was supported in part by the March of the Dimes Prematurity  
11  
12 371 Research Center at Stanford University, and Stanford Child Health Research Institute.  
13  
14  
15 372 The funders had no role in study design, data collection and analysis, decision to publish,  
16  
17 373 or preparation of the manuscript. Grant numbers are N/A.

18  
19  
20 374 **Conflict of Interest:** The authors report no conflict of interest.

21  
22  
23 375 **Author contributions:** XBL, KGS, and HJC contributed to concept development and  
24  
25 376 design.

26  
27  
28 377 JY, RJW, and DKS contributed to the acquisition of data.

29  
30  
31 378 KGS, SH, LZ, XY, LT, LM, SL, RJW, GMS, DKS, JCW and DBM contributed to the  
32  
33 379 analysis and interpretation of data.

34  
35  
36 380 KGS and SH drafted the manuscript.

37  
38  
39 381 JY, LZ, LT, XY, LM, SL, RJW, GMS, DKS, HJC, JCW, DBM, and XBL critically  
40  
41 382 revised the manuscript.

42  
43  
44 383 All the authors gave final approval of the version to be submitted and agreed to be  
45  
46 384 accountable for all aspects of the work.

47  
48  
49 385 **Data and materials availability:** The datasets used and/or analyzed in this study are  
50  
51 386 available upon request to the corresponding author.

52  
53  
54 387

**References**

1. Brownfoot FC, Gagliardi DI, Bain E, et al. Different corticosteroids and regimens for accelerating fetal lung maturation for women at risk of preterm birth. *Cochrane Database Syst Rev* 2013(8):CD006764. doi: 10.1002/14651858.CD006764.pub3
2. Raju TN, Mercer BM, Burchfield DJ, et al. Periviable birth: executive summary of a Joint Workshop by the Eunice Kennedy Shriver National Institute of Child Health and Human Development, Society for Maternal-Fetal Medicine, American Academy of Pediatrics, and American College of Obstetricians and Gynecologists. *Journal of perinatology : official journal of the California Perinatal Association* 2014;34(5):333-42. doi: 10.1038/jp.2014.70
3. Vohr B. Long-term outcomes of moderately preterm, late preterm, and early term infants. *Clin Perinatol* 2013;40(4):739-51. doi: 10.1016/j.clp.2013.07.006
4. Pereira AP, Dias MA, Bastos MH, et al. Determining gestational age for public health care users in Brazil: comparison of methods and algorithm creation. *BMC Res Notes* 2013;6:60. doi: 10.1186/1756-0500-6-60
5. Peek MJ, Devonald KJ, Beilby R, et al. The value of routine early pregnancy ultrasound in the antenatal booking clinic. *Aust N Z J Obstet Gynaecol* 1994;34(2):140-3.
6. Committee Opinion No 700: Methods for Estimating the Due Date. *Obstetrics and gynecology* 2017;129(5):e150-e54. doi: 10.1097/aog.0000000000002046 [published Online First: 2017/04/21]
7. Roman A, Saccone G, Dude CM, et al. Midtrimester transvaginal ultrasound cervical length screening for spontaneous preterm birth in diamniotic twin pregnancies according to chorionicity. *European journal of obstetrics, gynecology, and reproductive biology* 2018;229:57-63. doi: 10.1016/j.ejogrb.2018.08.006 [published Online First: 2018/08/15]
8. Erkamp JS, Voerman E, Steegers EAP, et al. Second and third trimester fetal ultrasound population screening for risks of preterm birth and small-size and large-size for gestational age at birth: a population-based prospective cohort study. *BMC Med* 2020;18(1):63. doi: 10.1186/s12916-020-01540-x [published Online First: 2020/04/08]
9. Dziadosz M, Bennett TA, Dolin C, et al. Uterocervical angle: a novel ultrasound screening tool to predict spontaneous preterm birth. *American journal of obstetrics and gynecology* 2016;215(3):376.e1-7. doi: 10.1016/j.ajog.2016.03.033 [published Online First: 2016/03/29]
10. Jehan I, Zaidi S, Rizvi S, et al. Dating gestational age by last menstrual period, symphysis-fundal height, and ultrasound in urban Pakistan. *International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics* 2010;110(3):231-4. doi: 10.1016/j.ijgo.2010.03.030
11. Wilson K, Hawken S, Potter BK, et al. Accurate prediction of gestational age using newborn screening analyte data. *American journal of obstetrics and gynecology* 2016;214(4):513 e1-13 e9. doi: 10.1016/j.ajog.2015.10.017



- 1  
2  
3 432 12. Knight AK, Craig JM, Theda C, et al. An epigenetic clock for gestational age at  
4 433 birth based on blood methylation data. *Genome biology* 2016;17(1):206. doi:  
5 434 10.1186/s13059-016-1068-z  
6  
7 435 13. Jelliffe-Pawlowski LL, Norton ME, Baer RJ, et al. Gestational dating by metabolic  
8 436 profile at birth: a California cohort study. *American journal of obstetrics and*  
9 437 *gynecology* 2016;214(4):511 e1-11 e13. doi: 10.1016/j.ajog.2015.11.029  
10 438 14. Aghaepour N, Ganio EA, McIlwain D, et al. An immune clock of human  
11 439 pregnancy. *Sci Immunol* 2017;2(15) doi: 10.1126/sciimmunol.aan2946  
12 440 15. Liang L, Rasmussen M-LH, Piening B, et al. Metabolic Dynamics and Prediction of  
13 441 Gestational Age and Time to Delivery in Pregnant Women. *Cell*  
14 442 2020;181(7):1680-92.e15. doi: <https://doi.org/10.1016/j.cell.2020.05.002>  
15 443 16. Ngo TTM, Moufarrej MN, Rasmussen MH, et al. Noninvasive blood tests for fetal  
16 444 development predict gestational age and preterm delivery. *Science*  
17 445 2018;360(6393):1133-36. doi: 10.1126/science.aar3819  
18 446 17. Aghaepour N, Lehallier B, Baca Q, et al. A proteomic clock of human pregnancy.  
19 447 *American journal of obstetrics and gynecology* 2018;218(3):347 e1-47 e14.  
20 448 doi: 10.1016/j.ajog.2017.12.208  
21 449 18. Esplin MS, Merrell K, Goldenberg R, et al. Proteomic identification of serum  
22 450 peptides predicting subsequent spontaneous preterm birth. *American journal of*  
23 451 *obstetrics and gynecology* 2011;204(5):391 e1-8. doi:  
24 452 10.1016/j.ajog.2010.09.021  
25 453 19. Saade GR, Boggess KA, Sullivan SA, et al. Development and validation of a  
26 454 spontaneous preterm delivery predictor in asymptomatic women. *American*  
27 455 *journal of obstetrics and gynecology* 2016;214(5):633 e1-33 e24. doi:  
28 456 10.1016/j.ajog.2016.02.001  
29 457 20. Peaceman AM, Andrews WW, Thorp JM, et al. Fetal fibronectin as a predictor of  
30 458 preterm birth in patients with symptoms: a multicenter trial. *American*  
31 459 *journal of obstetrics and gynecology* 1997;177(1):13-8.  
32 460 21. Strauss JF, 3rd, Romero R, Gomez-Lopez N, et al. Spontaneous preterm birth:  
33 461 advances toward the discovery of genetic predisposition. *American journal of*  
34 462 *obstetrics and gynecology* 2018;218(3):294-314 e2. doi:  
35 463 10.1016/j.ajog.2017.12.009  
36 464 22. Virgiliou C, Gika HG, Witting M, et al. Amniotic Fluid and Maternal Serum  
37 465 Metabolic Signatures in the Second Trimester Associated with Preterm  
38 466 Delivery. *Journal of proteome research* 2017;16(2):898-910. doi:  
39 467 10.1021/acs.jproteome.6b00845  
40 468 23. Li J, Lu YP, Reichetzeder C, et al. Maternal PCaaC38:6 is Associated With Preterm  
41 469 Birth - a Risk Factor for Early and Late Adverse Outcome of the Offspring.  
42 470 *Kidney Blood Press Res* 2016;41(3):250-7. doi: 10.1159/000443428  
43 471 24. Hawdon JM, Ward Platt Mp Fau - Aynsley-Green A, Aynsley-Green A. Patterns of  
44 472 metabolic adaptation for preterm and term infants in the first neonatal week.  
45 473 (1468-2044 (Electronic))  
46 474 25. Robinson HP. Sonar measurement of fetal crown-rump length as means of  
47 475 assessing maturity in first trimester of pregnancy. (0007-1447 (Print))  
48 476 26. Dunn WB, Broadhurst D, Begley P, et al. Procedures for large-scale metabolic  
49 477 profiling of serum and plasma using gas chromatography and liquid

- 1  
2  
3 478 chromatography coupled to mass spectrometry. *Nat Protoc* 2011;6(7):1060-  
4 479 83. doi: 10.1038/nprot.2011.335
- 5  
6 480 27. Sumner LW, Amberg A, Barrett D, et al. Proposed minimum reporting standards  
7 481 for chemical analysis Chemical Analysis Working Group (CAWG)  
8 482 Metabolomics Standards Initiative (MSI). *Metabolomics* 2007;3(3):211-21.  
9 483 doi: 10.1007/s11306-007-0082-2
- 10 484 28. Meis PJ, Klebanoff M, Thom E, et al. Prevention of recurrent preterm delivery by  
11 485 17 alpha-hydroxyprogesterone caproate. *The New England journal of*  
12 486 *medicine* 2003;348(24):2379-85. doi: 10.1056/NEJMoa035140
- 13  
14 487 29. Evenson KR, Siega-Riz AM, Savitz DA, et al. Vigorous leisure activity and  
15 488 pregnancy outcome. *Epidemiology* 2002;13(6):653-9. doi:  
16 489 10.1097/01.EDE.0000021463.45041.95
- 17 490 30. McIntire DD, Leveno KJ. Neonatal mortality and morbidity rates in late preterm  
18 491 births compared with births at term. *Obstetrics and gynecology*  
19 492 2008;111(1):35-41. doi: 10.1097/01.AOG.0000297311.33046.73
- 20  
21 493 31. Henderson-Smart DJ. The effect of gestational age on the incidence and duration  
22 494 of recurrent apnoea in newborn babies. *Aust Paediatr J* 1981;17(4):273-6.
- 23 495 32. Khashu M, Narayanan M, Bhargava S, et al. Perinatal outcomes associated with  
24 496 preterm birth at 33 to 36 weeks' gestation: a population-based cohort study.  
25 497 *Pediatrics* 2009;123(1):109-13. doi: 10.1542/peds.2007-3743
- 26  
27 498 33. Lee DS, Park J, Kay KA, et al. The implications of human metabolic network  
28 499 topology for disease comorbidity. *Proceedings of the National Academy of*  
29 500 *Sciences of the United States of America* 2008;105(29):9880-5. doi:  
30 501 10.1073/pnas.0802208105
- 31 502 34. Baumgartner C, Bohm C, Baumgartner D, et al. Supervised machine learning  
32 503 techniques for the classification of metabolic disorders in newborns.  
33 504 *Bioinformatics* 2004;20(17):2985-96. doi: 10.1093/bioinformatics/bth343
- 34  
35 505 35. Huang S, Chong N, Lewis NE, et al. Novel personalized pathway-based  
36 506 metabolomics models reveal key metabolic pathways for breast cancer  
37 507 diagnosis. *Genome medicine* 2016;8(1):34. doi: 10.1186/s13073-016-0289-9
- 38 508 36. Grootendorst-van Mil NH, Tiemeier H, Steenweg-de Graaff J, et al. Maternal  
39 509 plasma n-3 and n-6 polyunsaturated fatty acids during pregnancy and  
40 510 features of fetal health: Fetal growth velocity, birth weight and duration of  
41 511 pregnancy. *Clin Nutr* 2018;37(4):1367-74. doi: 10.1016/j.clnu.2017.06.010
- 42  
43 512 37. Menon R, Bonney EA, Condon J, et al. Novel concepts on pregnancy clocks and  
44 513 alarms: redundancy and synergy in human parturition. *Human reproduction*  
45 514 *update* 2016;22(5):535-60. doi: 10.1093/humupd/dmw022
- 46  
47 515 38. Gao L, Rabbitt EH, Condon JC, et al. Steroid receptor coactivators 1 and 2 mediate  
48 516 fetal-to-maternal signaling that initiates parturition. *The Journal of clinical*  
49 517 *investigation* 2015;125(7):2808-24. doi: 10.1172/JCI78544
- 50 518 39. Morillon A-C, Yakkundi S, Thomas G, et al. Association between phospholipid  
51 519 metabolism in plasma and spontaneous preterm birth: a discovery lipidomic  
52 520 analysis in the cork pregnancy cohort. *Metabolomics* 2020;16(2):19. doi:  
53 521 10.1007/s11306-020-1639-6
- 54  
55 522 40. Bennett PR, Rose MP, Myatt L, et al. Preterm labor: Stimulation of arachidonic  
56 523 acid metabolism in human amnion cells by bacterial products. *American*

1  
2  
3 524 *journal of obstetrics and gynecology* 1987;156(3):649-55. doi:  
4 525 [https://doi.org/10.1016/0002-9378\(87\)90070-6](https://doi.org/10.1016/0002-9378(87)90070-6)  
5  
6 526 41. Contreras MT, Gallardo MJ, Betancourt LR, et al. Correlation between plasma  
7 527 levels of arginine and citrulline in preterm and full-term neonates:  
8 528 Therapeutical implications. *Journal of Clinical Laboratory Analysis*  
9 529 2017;31(6):e22134. doi: 10.1002/jcla.22134  
10 530

## 13 531 **Figure Legends**

14  
15  
16 532 **Fig. 1.** Study design. Models were developed separately to estimate gestational age  
17  
18 533 during full-term pregnancy, and to predict the risk of preterm birth. Both models were  
19  
20 534 developed with the SU cohort and validated with the UAB cohort.

21  
22  
23 535 **Fig. 2.** Cohort construction. Each line represents an individual patient. Diamond and  
24  
25 536 triangle markers indicate sample collection dates and delivery dates, respectively. The red  
26  
27 537 dashed line represents 37 weeks' gestational age.

28  
29  
30 538 **Fig. 3.** The importance of the top 10 metabolic pathways in the gestational age estimation  
31  
32 539 model. Pathways either positively or negatively correlated gestational age.

33  
34  
35 540 **Fig. 4.** Gestational age estimates of the gestational age model with the SU ( $R^2=0.98$ ,  
36  
37 541 RMSE=1.09 weeks) and UAB cohorts ( $R^2 = 0.81$ , RMSE = 2.36 weeks).

38  
39  
40  
41 542 **Fig. 5.** (A) Univariate analysis of the 10 metabolic pathways in the preterm birth  
42  
43 543 prediction model. Odds ratio of each pathway was calculated.  $*P<0.05$ ,  $**P<0.01$ ,  
44  
45 544  $***P<0.005$ . (B) The importance of the metabolic pathways in the preterm birth  
46  
47 545 prediction model. Pathways were either up- or down-regulated in relation to preterm birth.

48  
49  
50 546 **Fig. 6.** (A) Prediction of preterm birth risk grouped by full-term and preterm birth  
51  
52 547 patients (top) and over the course of gestation (bottom). (B) AUC performance of the

1  
2  
3 548 prediction in SU and UAB cohorts.  $P$  was calculated using Mann–Whitney U test. wks:  
4  
5 549 weeks' gestational age.  
6  
7

8 550 **Fig. 7.** Performance of the preterm birth prediction model. (A) A contingency table  
9  
10 551 showing the number of samples in each category. (B) Sensitivity, specificity, PPV, and  
11  
12 552 NPV together with the 95% confidence intervals.  
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

For peer review only

1  
2  
3 **553 Appendix Captions**  
4

5  
6 **554 Fig. A.1** False discovery rate (FDR) analysis of the metabolic pathways significantly  
7  
8 **555** associated with the GA in full-term pregnancies. Pearson  $|r|$  was calculated as the  
9  
10 **556** correlation between metabolite serological abundance and GA. Only the metabolites with  
11  
12 **557** a Pearson  $|r|$  higher than the threshold would be selected as part of the significant  
13  
14 **558** pathways. FDR was estimated by a permutation-based method (permutation N=1000).  
15  
16

17  
18 **559 Fig. A.2** Profile of the metabolic pathways in the GA estimation model over the course of  
19  
20 **560** gestation on SU cohort. All pathways are (A) positively or (B) negatively correlated to  
21  
22 **561** the GA (FDR<1%). Profile of each pathway was calculated as the weighted sum of the z-  
23  
24 **562** score normalized metabolite serological abundances divided by the number of  
25  
26 **563** metabolites. Means  $\pm$  standard errors at each time point were plotted.  
27  
28

29  
30 **564 Fig. A.3** Univariate analysis of the 33 metabolic pathways in the GA estimation model.  
31  
32 **565** Pearson correlation coefficient of each pathway to GA was calculated. \* $P<0.05$ ,  
33  
34 **566** \*\* $P<0.01$ , \*\*\* $P<0.005$ .  
35  
36

37 **567 Fig. A.4** Comparison of GA estimates using the model and US measurements. (A)  
38  
39 **568** Distributions of differences between GA measured by US and GA estimated by the  
40  
41 **569** model, in T2 (weeks 14–27), T3 (weeks 28–40), and T2+T3. n represents the number of  
42  
43 **570** full-term patients included. (B) Error distribution of GA estimation on a combination of  
44  
45 **571** SU and UAB cohorts in T2, T3, and T2+T3.  
46  
47

48  
49 **572 Fig. A.5** False discovery rate (FDR) analysis of the metabolic pathways significantly  
50  
51 **573** associated with PTB. Mann-Whitney U test  $P$  measured the difference in metabolite  
52  
53 **574** serological abundances between full-term pregnancies and pregnancies ending in PTB.  
54  
55  
56  
57  
58  
59

1  
2  
3 575 Only metabolites with a Mann-Whitney U test  $P$  lower than the threshold were selected  
4  
5 576 as part of the significant pathways. FDR was estimated by a permutation-based method  
6  
7  
8 577 (permutation  $N=1000$ ).  
9

10 578 **Fig. A.6** Stratification of patients by the classification model prediction on the UAB  
11  
12 579 cohort. PPV was corrected by bootstrapping the full-term patients to reach the population  
13  
14 580 PTB prevalence of 12.5% on singleton births. Two horizontal dashed lines represent the  
15  
16 581 population mean of PTB risk that is 12.5% (black) and the PPV (= 0.70; red) at the high-  
17  
18 582 risk cutoff. The grey dashed line indicates the high-risk cutoff value (= 0.52). The grey  
19  
20 583 area represents the 95% confidence interval of the PPV. The box plot at the bottom shows  
21  
22 584 the classification model value distribution stratified by the samples. GAB: GA at birth.  
23  
24 585 wks: weeks of gestation.  
25  
26  
27  
28

29 586 **Fig. A.7** The performance of the IBP4/SHBG predictor and the metabolic model. The  
30  
31 587 results are stratified by the GA intervals with a BMI at 22–37 kg/m<sup>2</sup> (A), and by BMI  
32  
33 588 values with a GA interval of 5–20 weeks (B).  
34  
35  
36

37 589 **Fig. A.8** (A) False discovery rate (FDR) analysis of the metabolites and metabolic  
38  
39 590 pathways significantly associated with GA in full-term pregnancies. Pearson  $|r|$  was  
40  
41 591 calculated as the correlation between metabolite serological abundance and GA. Only the  
42  
43 592 metabolites with a Pearson  $|r|$  higher than the threshold (=0.35) would be selected as part  
44  
45 593 of the significant pathways. FDR was estimated by a permutation-based method  
46  
47 594 (permutation  $N=1000$ ). (B) A comparison of RMSE of the GA estimation model trained  
48  
49 595 by pathways and the model trained by metabolites. All metabolites had a Pearson  $|r|>0.35$ .  
50  
51 596 RMSE was measured with the full-term samples of the validation (UAB) cohort.  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 597 **Fig. A.9** (A) False discovery rate (FDR) analysis of the metabolites and metabolic  
4  
5 598 pathways significantly associated with the PTB. Mann-Whitney U test  $P$  measured the  
6  
7 599 difference in metabolite serological abundances between full-term pregnancies and  
8  
9 600 pregnancies ending in PTB. Only the metabolites with a Mann-Whitney U test  $P$  lower  
10  
11 601 than the threshold ( $=0.05$ ) would be selected as part of the significant pathways. FDR was  
12  
13 602 estimated by a permutation-based method (permutation  $N=1000$ ). (B) A comparison of  
14  
15 603 the AUC of the PTB classification model utilizing pathways and the model utilizing  
16  
17 604 metabolites. All the metabolites had a Mann-Whitney U test  $P < 0.05$ . AUC was  
18  
19 605 measured with the samples of the validation (UAB) cohort.  
20  
21  
22  
23

24 606 **Table A.1** Sensitivity and specificity of the XGBoost model with respect to the cutoff  
25  
26 607 point.  
27  
28

29 608 **Text A.1** Metabolic compound selection, pathway computation, and model development  
30  
31

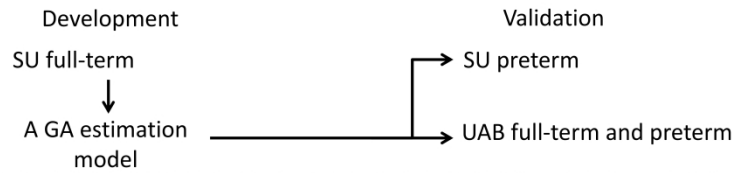
32 609 **Text A.2** Metabolite model vs. IBP4/SHBG in predicting PTB  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

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

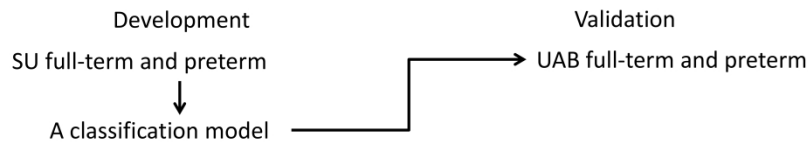
**A. Study cohort**

SU Cohort	UAB Cohort
20 full-term, 16 preterm	9 full-term, 13 preterm

**B. To estimate GA for full-term pregnancies**



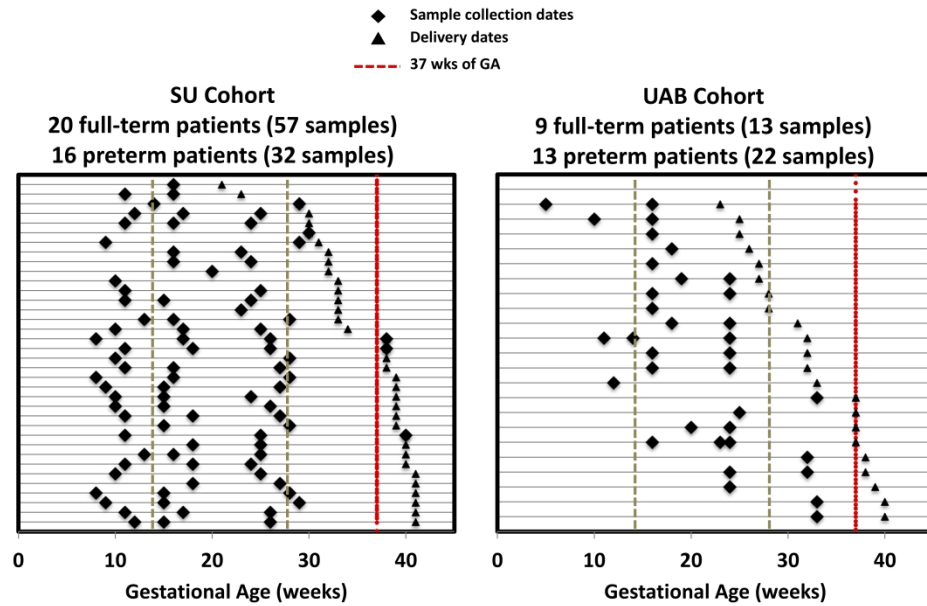
**C. To identify women at risk of PTB**



Study design. Models were developed separately to estimate gestational age during full-term pregnancy, and to predict the risk of preterm birth. Both models were developed with the SU cohort and validated with the UAB cohort.

254x190mm (600 x 600 DPI)

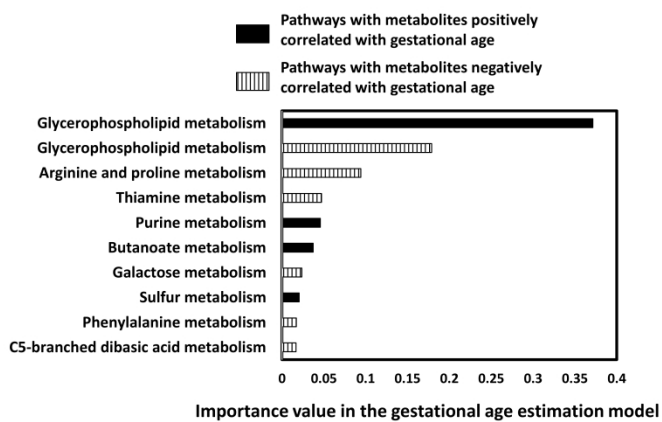




31 Cohort construction. Each line represents an individual patient. Diamond and triangle markers indicate  
32 sample collection dates and delivery dates, respectively. The red dashed line represents 37 weeks'  
33 gestational age.

34 254x190mm (600 x 600 DPI)

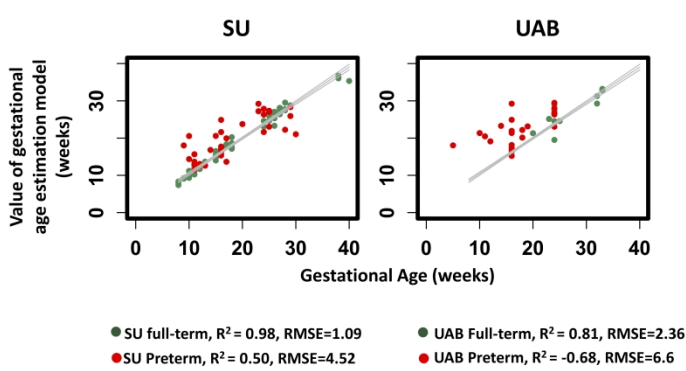
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



The importance of the top 10 metabolic pathways in the gestational age estimation model. Pathways either positively or negatively correlated gestational age.

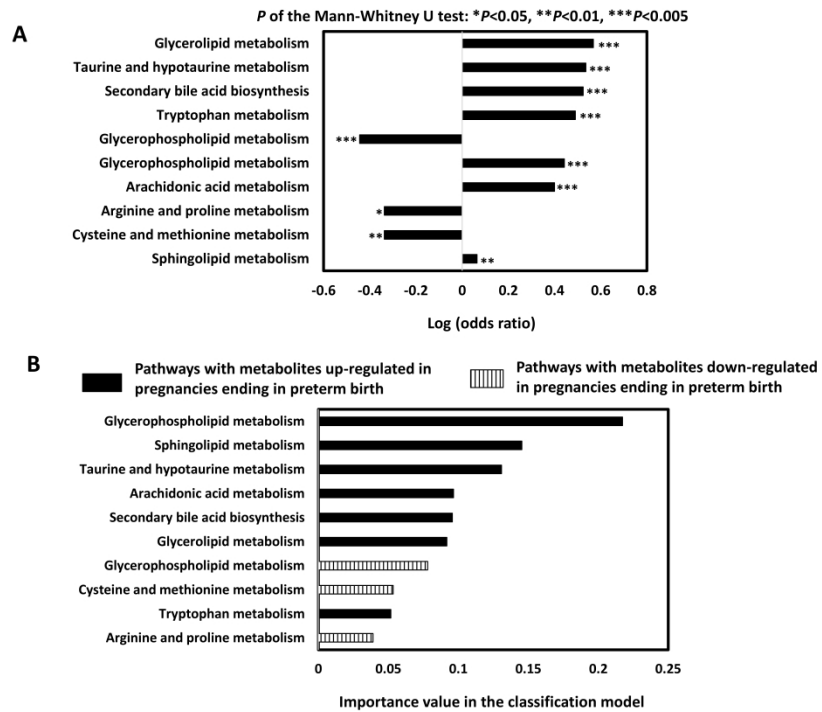
254x190mm (600 x 600 DPI)

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



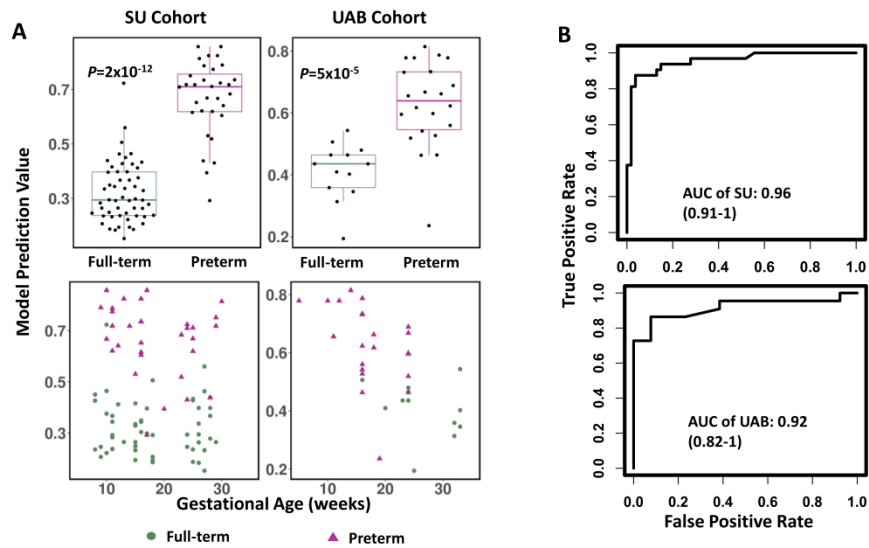
Gestational age estimates of the gestational age model with the SU ( $R^2=0.98$ , RMSE=1.09 weeks) and UAB cohorts ( $R^2 = 0.81$ , RMSE = 2.36 weeks).

254x190mm (600 x 600 DPI)



(A) Univariate analysis of the 10 metabolic pathways in the preterm birth prediction model. Odds ratio of each pathway was calculated. \**P*<0.05, \*\**P*<0.01, \*\*\**P*<0.005. (B) The importance of the metabolic pathways in the preterm birth prediction model. Pathways were either up- or down-regulated in relation to preterm birth.

254x190mm (600 x 600 DPI)



(A) Prediction of preterm birth risk grouped by full-term and preterm birth patients (top) and over the course of gestation (bottom). (B) AUC performance of the prediction in SU and UAB cohorts. P was calculated using Mann-Whitney U test. wks: weeks' gestational age.

254x190mm (600 x 600 DPI)

**A**

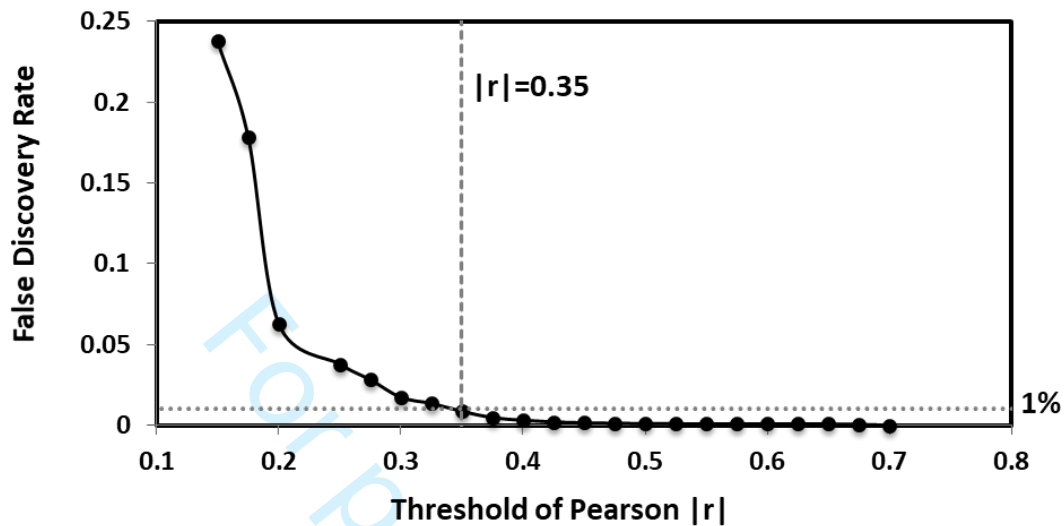
	<b>SU</b>			<b>UAB</b>	
	Preterm	Full-term		Preterm	Full-term
Classified as Preterm	28	2	Classified as Preterm	19	1
Classified as Full-term	4	52	Classified as Full-term	3	12

**B**

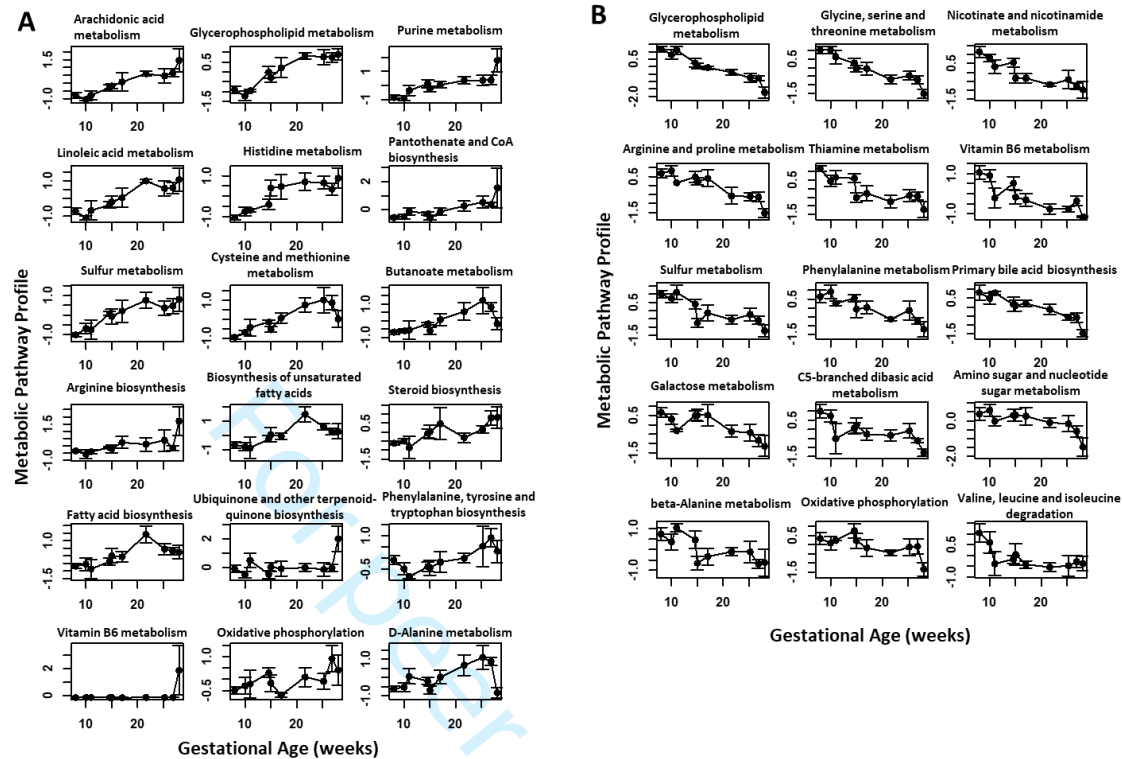
Cohort	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
<b>SU</b>	0.88 (0.71-0.97)	0.96 (0.87-1)	0.93 (0.78-0.99)	0.93 (0.83-0.98)
<b>UAB</b>	0.86 (0.65-0.97)	0.92 (0.64-1)	0.95 (0.75-1)	0.80 (0.52-0.96)

Performance of the preterm birth prediction model. (A) A contingency table showing the number of samples in each category. (B) Sensitivity, specificity, PPV, and NPV together with the 95% confidence intervals.

254x190mm (600 x 600 DPI)

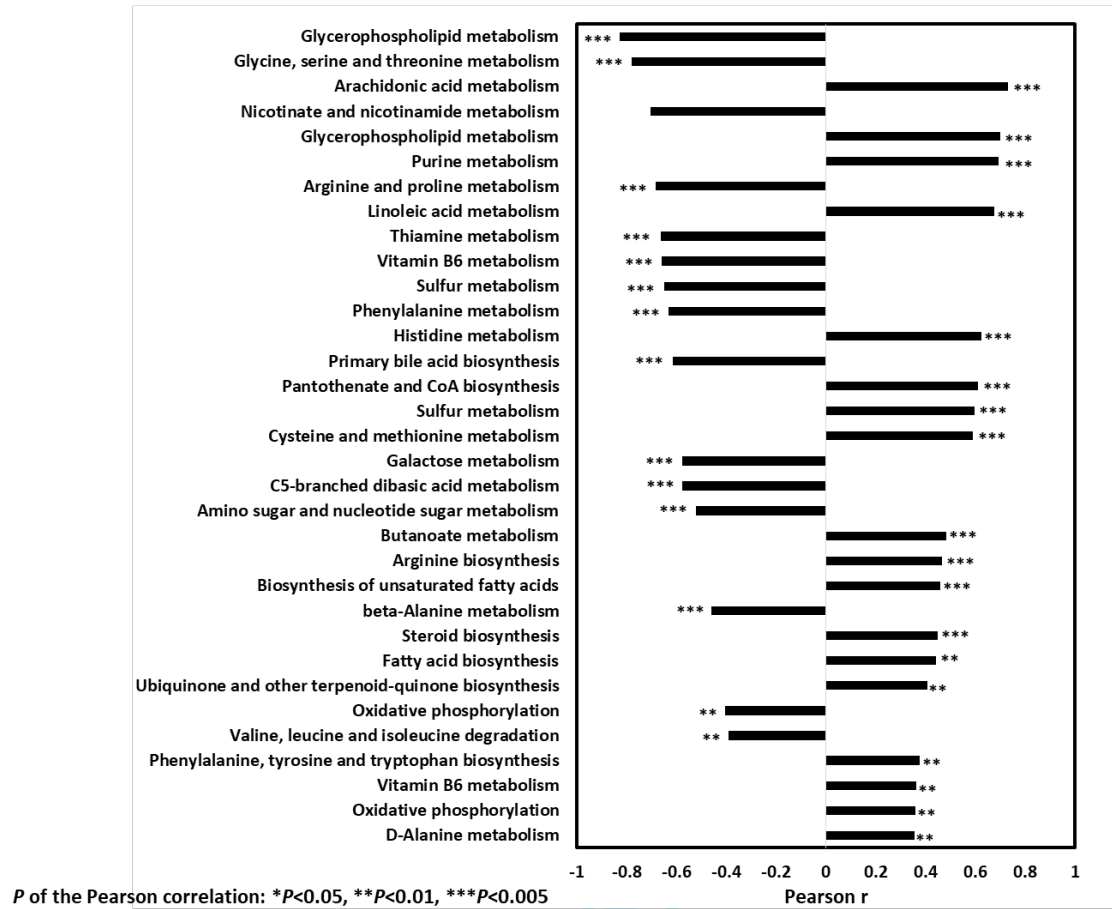


**Fig. A.1.** False discovery rate (FDR) analysis of the metabolic pathways significantly associated with the GA in full-term pregnancies. Pearson  $|r|$  was calculated as the correlation between metabolite serological abundance and GA. Only the metabolites with a Pearson  $|r|$  higher than the threshold would be selected as part of the significant pathways. FDR was estimated by a permutation-based method (permutation  $N=1000$ ).



**Fig. A.2.** Profile of the metabolic pathways in the GA estimation model over the course of gestation on SU cohort. All pathways are (A) positively or (B) negatively correlated to the GA (FDR<1%). Profile of each pathway was calculated as the weighted sum of the z-score normalized metabolite serological abundances divided by the number of metabolites. Mean  $\pm$  standard error of the mean at each time point was plotted.

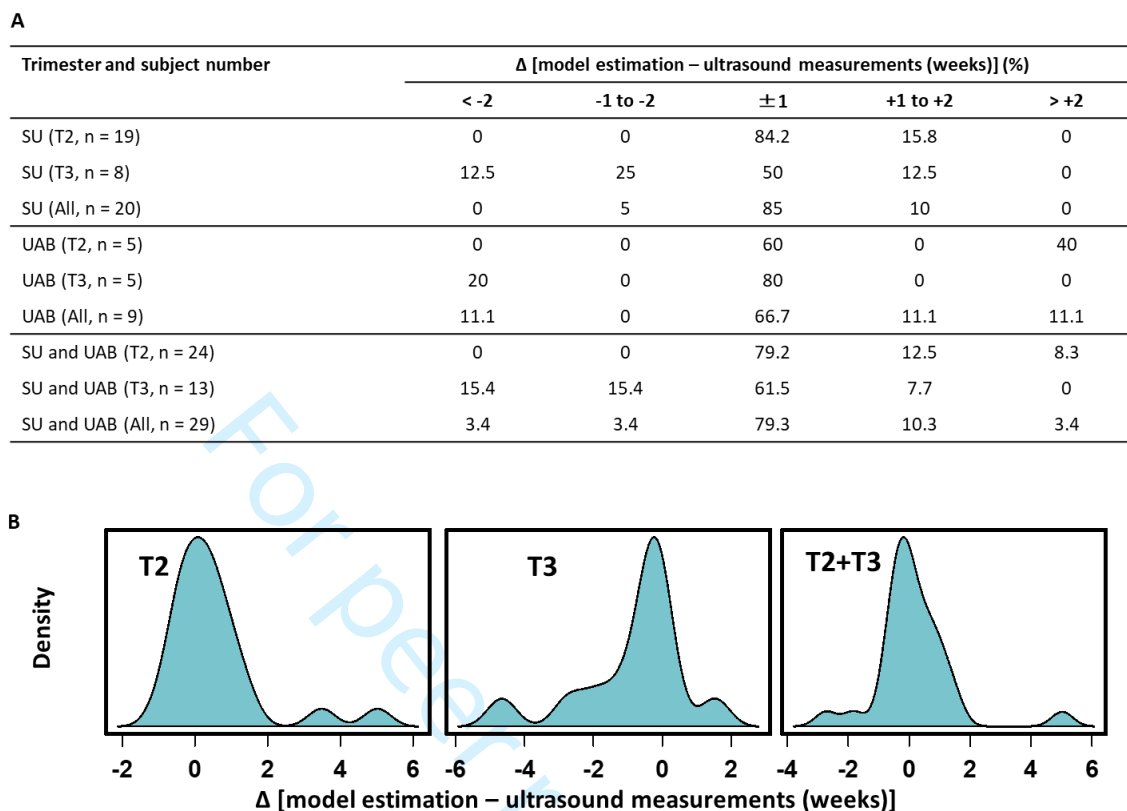




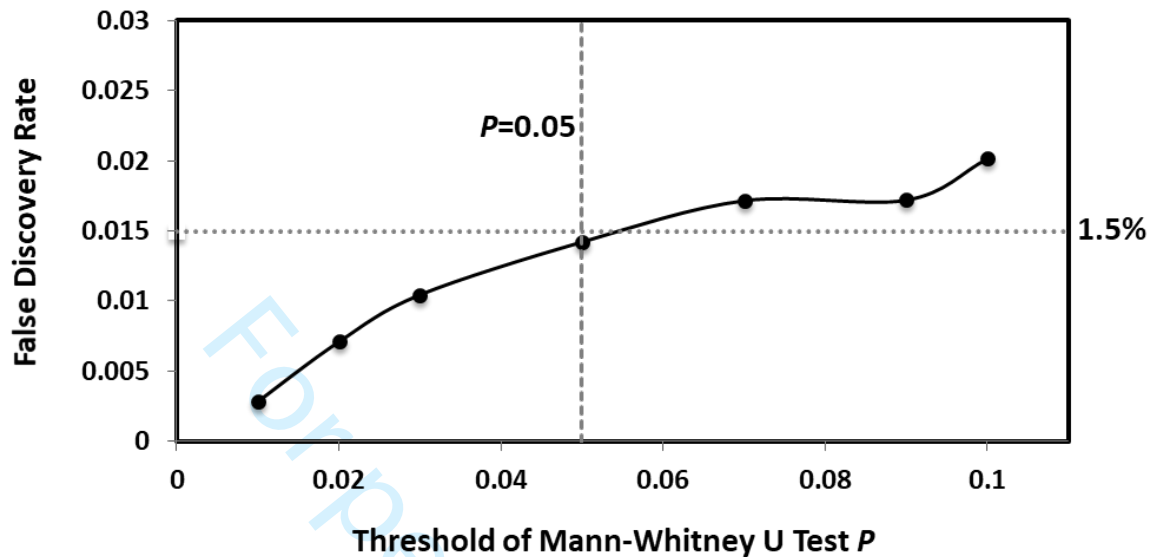
**Fig. A.3.** Univariate analysis of the 33 metabolic pathways in the GA estimation model.

Pearson correlation coefficient  $r$  of each pathway to GA was calculated. \* $P<0.05$ ,

\*\* $P<0.01$ , \*\*\* $P<0.005$ .

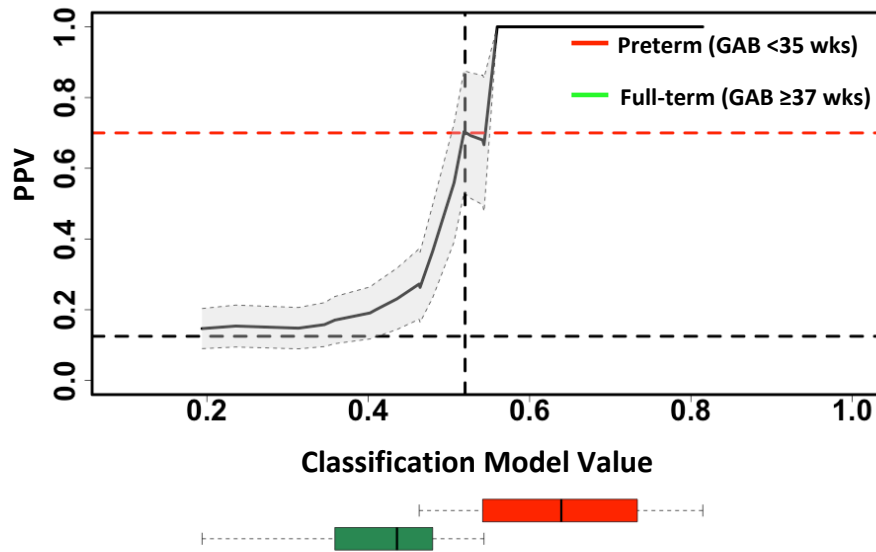


**Fig. A.4.** Comparison of GA estimates using the model and US measurements. (A) Distributions of differences between GA measured by US and GA estimated by the model, in T2 (weeks 14–27), T3 (weeks 28–40), and T2+T3. n represents the number of full-term patients included. (B) Error distribution of GA estimation on a combination of SU and UAB cohorts in T2, T3, and T2+T3.

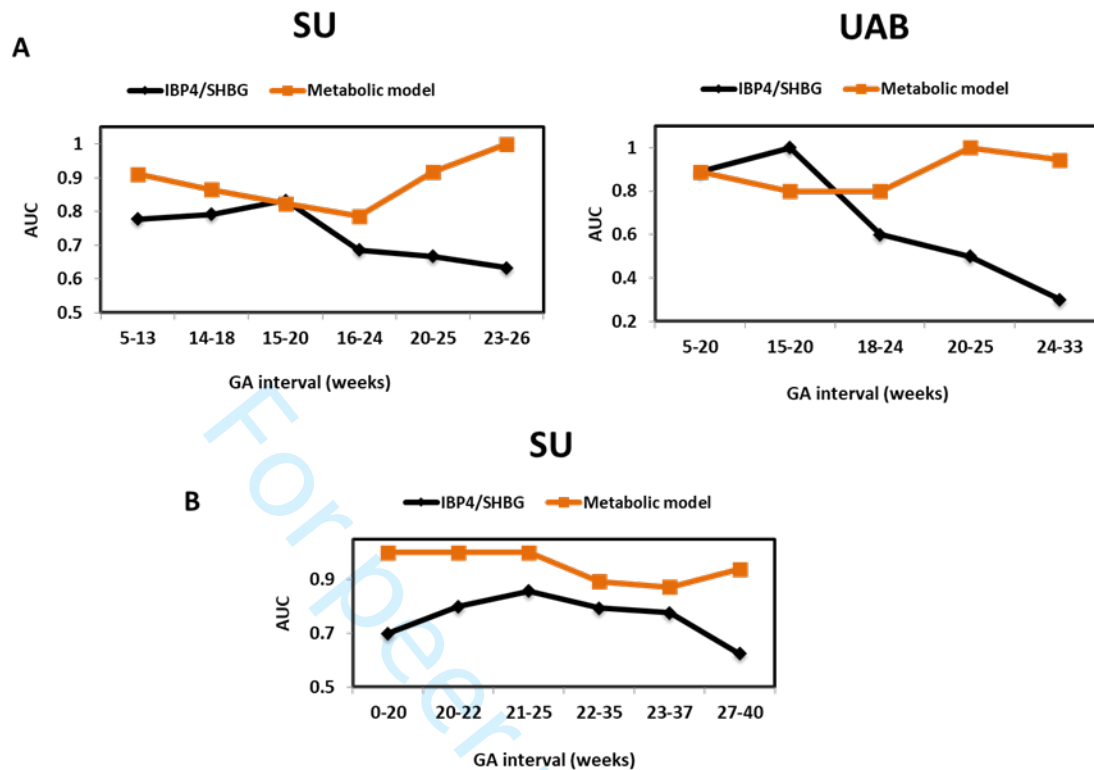


**Fig. A.5.** False discovery rate (FDR) analysis of the metabolic pathways significantly associated with PTB. Mann-Whitney U test  $P$  measured the difference in metabolite serological abundances between full-term pregnancies and pregnancies ending in PTB. Only metabolites with a Mann-Whitney U test  $P$  lower than the threshold were selected as part of the significant pathways. FDR was estimated by a permutation-based method (permutation  $N=1000$ ).

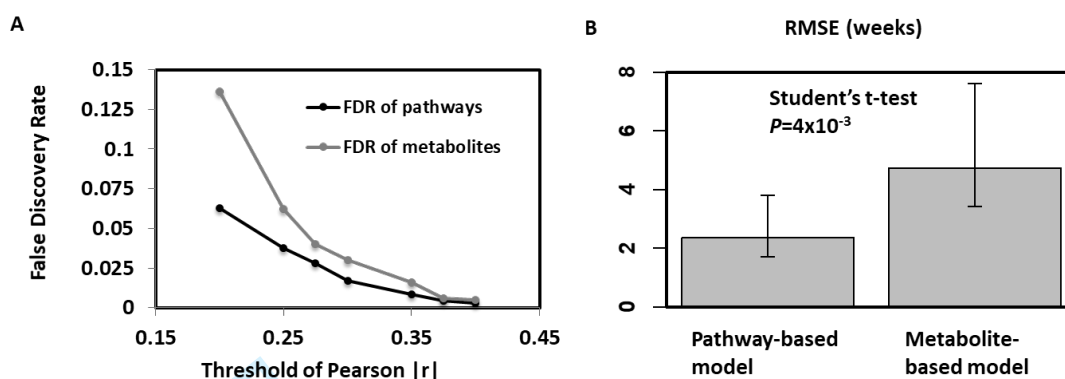
Population-corrected PPV: 0.70. which is **5.6** times higher than the general population risk in Alabama (12.5%)



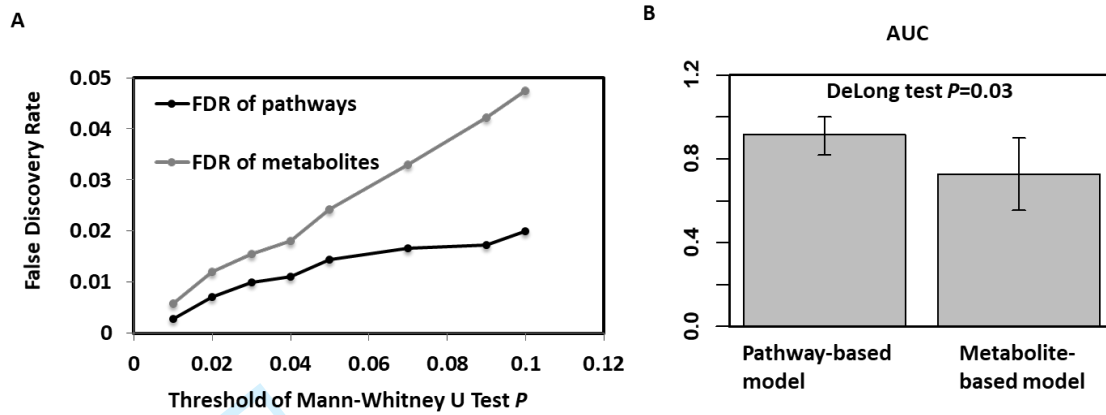
**Fig. A.6.** Stratification of patients by the classification model prediction on the UAB cohort. PPV was corrected by bootstrapping the full-term patients to reach the population PTB prevalence of 12.5% on singleton births in Alabama. Two horizontal dashed lines represent the population mean of PTB risk that is 12.5% (black) and the PPV (= 0.70; red) at the high-risk cutoff. The grey dashed line indicates the high-risk cutoff value (= 0.52). The grey area represents the 95% confidence interval of the PPV. The box plot at the bottom shows the classification model value distribution stratified by the samples. GAB: gestational age at birth. wks: weeks' GA.



**Fig. A.7.** The performance of the IBP4/SHBG predictor and the metabolic model. The results are stratified by the GA intervals with a BMI at 22–37 kg/m<sup>2</sup> (A), and by BMI values with a GA interval of 5–20 weeks (B).



**Fig. A.8.** (A) False discovery rate (FDR) analysis of the metabolites and metabolic pathways significantly associated with the GA in full-term pregnancies. Pearson  $|r|$  was calculated as the correlation between metabolite serological abundance and GA. Only the metabolites with a Pearson  $|r|$  higher than the threshold ( $=0.35$ ) would be selected as part of the significant pathways. FDR was estimated by a permutation-based method (permutation  $N=1000$ ). (B) A comparison of RMSE of the GA estimation model trained by pathways and the model trained by metabolites. All metabolites had a Pearson  $|r|>0.35$ . RMSE was measured with the full-term samples of the validation (UAB) cohort.



**Fig. A.9.** (A) False discovery rate (FDR) analysis of the metabolites and metabolic pathways significantly associated with the PTB. Mann-Whitney U test  $P$  measured the difference in metabolite serological abundances between full-term pregnancies and pregnancies ending in PTB. Only the metabolites with a Mann-Whitney U test  $P$  lower than the threshold ( $=0.05$ ) would be selected as part of the significant pathways. FDR was estimated by a permutation-based method (permutation  $N=1000$ ). (B) A comparison of the AUC of the preterm birth classification model utilizing pathways and the model utilizing metabolites. All the metabolites had a Mann-Whitney U test  $P < 0.05$ . AUC was measured with the samples of the validation (UAB) cohort.

**Table A.1.** Sensitivity and specificity of the XGBoost model with respect to the cutoff point.

Cutoff	Cohort	Sensitivity	Specificity	Number of preterm samples identified by the model
<b>0.4</b>	<b>SU</b>	0.94	0.78	30
	<b>UAB</b>	0.95	0.31	21
<b>0.5</b>	<b>SU</b>	0.88	0.94	28
	<b>UAB</b>	0.86	0.85	19
<b>0.6</b>	<b>SU</b>	0.81	0.98	26
	<b>UAB</b>	0.59	1	13
<b>0.7</b>	<b>SU</b>	0.53	0.98	17
	<b>UAB</b>	0.32	1	7



## **Text A.1 Metabolic compound selection, pathway computation, and model development**

### *GA estimation*

Metabolites measured by targeted and untargeted MS were aggregated and filtered using Pearson correlation coefficient analyses in relation to GA. The remaining metabolites were mapped to pathways. The value of each pathway was calculated as the weighted sum of the normalized concentrations of metabolites on the pathway divided by the number of metabolites. The weight of each metabolite was the absolute value of the Pearson correlation coefficient in relation to GA. Metabolites having positive or negative coefficients were aggregated separately. That is, a pathway could have two values, one for metabolites positively correlated to GA, and the other for those negatively correlated to GA.

A supervised, cross-validated machine-learning technique XGBoost was developed with the pathway values of samples from full-term patients in the SU cohort. An ensemble of regression trees was generated to give a score estimating the GA. The model was validated on the UAB cohort. For a patient that had multiple samples, an 'integrated' GA estimate was calculated by shifting the GA estimates of every sample to a reference point for obtaining the median. Error distribution of GA estimation based on patients was calculated as the distribution of the differences between the 'integrated' GA estimates and the US measurement.

### *PTB prediction*

1  
2  
3 Samples collected before 35 weeks' GA were selected to build the model to predict PTB.  
4  
5 Mann–Whitney U test was used to select the initial candidate metabolites that were then  
6  
7 mapped to pathways. The value of each pathway was calculated as the weighted sum of  
8  
9 the normalized concentrations of metabolites on the pathway divided by the number of  
10  
11 metabolites. The weight of each metabolite was the absolute value of the ratio of median  
12  
13 of full-term samples to PTB samples. Like the GA estimation, pathways could have two  
14  
15 values that depended on the ratio of median greater or less than 1. An XGBoost model  
16  
17 was developed utilizing samples from the SU cohort and validated with the UAB cohort.  
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

## Text A.2 Metabolite model vs. IBP4/SHBG in predicting PTB

We conducted ELISA tests on the SU and UAB cohorts to evaluate the IBP4/SHBG signature, a predictor that was validated in a prospective study as a predictor of spontaneous PTB. Commercial kits Human IGFBP4 ELISA Kit (Abcam, Burlingame, CA, USA) and Human SHBG Quantikine ELISA Kit (R&D System Inc.) were used. AUC of the predictor was calculated in different GA intervals and with different maternal BMI values, and was compared to the performance of the metabolic model.

With a BMI of  $>22$  and  $\leq 37$  kg/m<sup>2</sup>, the AUC values of the IBP4/SHBG predictor peaked at 15–20 weeks' GA (SU: 0.833; UAB: 1), and dropped rapidly after 20 weeks (Figure A below). The AUC values were lower with extreme BMI (0.7 at BMI  $\leq 20$  kg/m<sup>2</sup> and 0.63 at BMI  $>27$  kg/m<sup>2</sup>; see Figure B below). These findings are consistent with the previous validation study. Compared with the IBP4/SHBG predictor, the metabolic model has a more stable AUC performance over the gestation and different BMI values in SU ( $P = 0.03$ ). In UAB at  $>18$  weeks' GA, the AUC of IBP4/SHBG dropped from 0.6 to 0.3, while the AUC of the metabolic model was above 0.8.

## STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	Page No
<b>Title and abstract</b>	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	1
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	2-3
<b>Introduction</b>			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	5
Objectives	3	State specific objectives, including any prespecified hypotheses	6
<b>Methods</b>			
Study design	4	Present key elements of study design early in the paper	6
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6,7
Participants	6	(a) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up <i>Case-control study</i> —Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls <i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants	6,7
		(b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed <i>Case-control study</i> —For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6,7,8
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	7
Bias	9	Describe any efforts to address potential sources of bias	8
Study size	10	Explain how the study size was arrived at	
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	7,8
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	8
		(b) Describe any methods used to examine subgroups and interactions	
		(c) Explain how missing data were addressed	
		(d) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed <i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy	
		(e) Describe any sensitivity analyses	

Continued on next page

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

<b>Results</b>			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	9
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	9
		(b) Indicate number of participants with missing data for each variable of interest	
		(c) <i>Cohort study</i> —Summarise follow-up time (eg, average and total amount)	9
Outcome data	15*	<i>Cohort study</i> —Report numbers of outcome events or summary measures over time	10
		<i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure	
		<i>Cross-sectional study</i> —Report numbers of outcome events or summary measures	
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	11-12
		(b) Report category boundaries when continuous variables were categorized	11-12
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	12
<b>Discussion</b>			
Key results	18	Summarise key results with reference to study objectives	13
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	16
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	14-15
Generalisability	21	Discuss the generalisability (external validity) of the study results	16
<b>Other information</b>			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	18

\*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

**Note:** An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at [www.strobe-statement.org](http://www.strobe-statement.org).