Vaginal bacterial load in the second trimester is associated with early preterm birth recurrence: a nested case-control study

Authors:
Laura Goodfellow,[1*] Marijn C. Verwijs,[2] Angharad Care,[1] Andrew Sharp,[1] Jelena Ivandic,[1] Borna Poljak,[3] Devender Roberts,[3] Christina Bronowski,[2] A. Christina Gill,[2] Alistair C. Darby,[4] Ana Alfirevic,[1] Bertram Muller-Myhsok,[1,5] Zarko Alfirevic,[1†] Janneke H.H.M. van de Wijgert[2,6†]

Affiliations:
1. Harris Wellbeing Research Centre, Department of Women's and Children's Health University of Liverpool, United Kingdom
2. Department of Clinical Infection, Microbiology and Immunology, Institute of Infection, Veterinary, and Ecological Sciences, University of Liverpool, Liverpool, United Kingdom.
3. Liverpool Women's NHS Foundation Trust, Liverpool, United Kingdom.
4. Centre for Genomic Research, University of Liverpool, Liverpool.
5. Department of Translational Research in Psychiatry, Max Planck Institute of Psychiatry, Munich, Germany.
6. Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

*Corresponding author. Address: University Department, Liverpool Women's Hospital, Crown Street, Liverpool, L8 7SS. Tel: 0151 7959557. Fax: 0151 795 9599 email: l.goodfellow@liverpool.ac.uk

†Contributed equally

Running title (60 characters): Vaginal microbiota & recurrent early preterm birth

Word Counts:
Abstract: 249 words (max. 250)
Introduction: 371 words (max. 400)
Text: 3,491 words (max. 3,500)
Tables/figures: 3 tables/1 figure (max. 4)
Appendices: 2

Abstract

Objective: To assess the association between vaginal microbiome (VMB) composition and recurrent early spontaneous preterm birth (sPTB)/preterm prelabour rupture of membranes (PPROM).

Design: Nested case-control study.

Setting: UK tertiary referral hospital.

Sample: High-risk women with previous sPTB/PPROM <34^{+0} weeks gestation who had a recurrence (n=22) or delivered at ≥37^{+0} weeks without PPROM (n=87).

Methods: Vaginal swabs collected between 15-22 weeks gestation were analysed by 16S rRNA gene sequencing and 16S quantitative PCR.

Main outcome measure: Recurrent early sPTB/PPROM.

Results: 28/109 high-risk women had anaerobic vaginal dysbiosis, with the remainder dominated by lactobacilli (*L. iners* 36/109, *L. crispatus* 23/109, or other 22/109). VMB type, diversity, and stability were not associated with recurrence. Women with a recurrence, compared to those without, had a higher median vaginal bacterial load (8.64 vs. 7.89 $\log_{10}$ cells/µl, adjusted odds ratio (aOR)=1.90, 95% confidence interval (CI)=1.01-3.56, p=0.047) and estimated *Lactobacillus* concentration (8.59 vs. 7.48 $\log_{10}$ cells/µl, aOR=2.35, CI=1.20-4.61, p=0.013). A higher recurrence risk was associated with higher median bacterial loads for each VMB type after stratification, although statistical significance was reached only for *L. iners*-domination (aOR=3.44, CI=1.06-11.15, p=0.040). Women with anaerobic dysbiosis or *L. iners*-domination had a higher median vaginal bacterial load than women with a VMB dominated by *L. crispatus* or other lactobacilli (8.54, 7.96, 7.63, and 7.53 $\log_{10}$ cells/µl, respectively).

Conclusions: Vaginal bacterial load is associated with early sPTB/PPROM recurrence. Domination by lactobacilli other than *L. iners* may protect women from developing high bacterial loads. Future PTB studies should quantify vaginal bacteria and yeasts.

Funding: Wellbeing of Women, London, UK

Keywords: vaginal microbiome; *Lactobacillus*; spontaneous preterm birth; preterm premature rupture of membranes.

Tweetable abstract: Increased vaginal bacterial load in the second trimester may be associated with recurrent early spontaneous preterm birth.

Introduction

Previous spontaneous preterm birth (sPTB) or preterm prelabour rupture of membranes (PPROM) is the strongest risk factor for recurrent sPTB/PPROM.[1] This may implicate a genetic contribution[2] but does not rule out a contribution of persistent microbes and/or inflammation in the female reproductive tract.[3]

16S ribosomal RNA (16S rRNA) sequencing was first used to assess the relationship between the bacterial vaginal microbiota (VMB) and PTB in 2014.[4] Since then over 2000 pregnant women have been assessed in 12 studies (Appendix A, Table A1) . The majority of these studies suggested a contribution of the VMB composition to PTB, but results have been inconsistent. The healthy vagina in pregnancy is dominated by *Lactobacillus* species but contains up to 80 different *Lactobacillus* and other bacterial taxa, and the optimal ways of summarising and categorising sequencing data to determine clinically relevant associations with PTB are yet to be determined.[5,6] Studies to date have shown that increased relative abundances of *L. crispatus, L. gasseri*, and *L. jensenii* were either not associated with PTB or associated with a reduced PTB risk, and increased relative abundances of bacterial vaginosis (BV)-associated anaerobes were either not associated with PTB or associated with an increased PTB risk (Table A1). Results for *L. iners* relative abundance were inconsistent, with 6 studies showing no association, two studies an increased risk, and one study a decreased risk (Table A1). Importantly, all studies except two[7,8] relied on compositional data, meaning that an increase in the proportion of one taxon always leads to a decrease in the proportions of other taxa.[9] Furthermore, an increase in a proportion does not necessarily indicate an increase in absolute quantity. VMB researchers have only recently begun to semi-quantify 16S rRNA sequencing data by combining sequencing data with a quantitative PCR (qPCR) of

the 16S rRNA gene.[10] The two studies to date that included some quantification reported opposite results, with one reporting a positive association between overall vaginal bacterial load and PTB risk,[7] and the other a negative association.[8]

Our aim was to assess associations between VMB compositions and recurrent early sPTB/PPROM (delivery at <34^{+0} weeks compared to delivery at ≥37^{+0} weeks) among high-risk pregnant women who had had a previous early sPTB/PPROM, using compositional and semi-quantitative 16S sequencing data.

Methods

Study design and population

Two cohorts of women at high and low-risk of PTB were enrolled at Liverpool Women's Hospital, UK, from 1 June 2015 until 31 December 2017. The hospital's patient and public engagement group helped guide the research and recruitment (Appendix A). High-risk was defined as a history of sPTB or PPROM at 16^{+0}-33^{+6} weeks gestation, and low-risk as having had at least one previous term birth (≥37^{+0} weeks) without PPROM and never having had a PTB or PPROM. In the low-risk cohort, but not in the high-risk cohort, women were excluded if they had had significant cervical surgery or previous obstetric or medical problems (Appendix A). For this nested case-control study, we selected 109 women from the high-risk cohort (N=137) and 145 women from the low-risk cohort (N=223) based on the pregnancy outcome (sPTB or PPROM 16^{+0}-33^{+6} weeks or birth ≥37^{+0} weeks in the high-risk cohort and birth ≥39^{+0} weeks in the low-risk cohort) and availability of valid VMB data (Figure 1). Birth outcomes were classified independently by two obstetricians experienced in PTB

113 management. When there was a discrepancy between them, the case was reviewed by a
114 third experienced obstetrician until the team reached consensus. Participants with medically
115 indicated PTBs, or for whom it was unclear whether a birth was truly spontaneous, were
116 excluded. We used relevant parts of the CROWN initiative core outcome set for PTB
117 research[11] (Appendix A).
118
119 *Clinical procedures*
120 Participants were invited to two study visits at approximately 16 weeks (15^{+1}-18^{+6} weeks)
121 and 20 weeks (19^{+0}-23^{+0}) gestation. At each visit, a study obstetrician interviewed the
122 participant; collected high vaginal swabs during a speculum examination for VMB
123 assessments at the University of Liverpool Center for Genomic Research (HydraFlock
124 standard tapered swabs, Medical Wire and Equipment, Corsham, UK) and for quantitative
125 assessment of foetal fibronectin and culture by the local NHS laboratory; determined
126 cervical length; and did a 3D foetal ultrasound. All other procedures during pregnancy, and
127 all treatments, were in accordance with UK national guidelines.[12] Preterm birth prevention
128 therapy (cerclage, vaginal pessary, or vaginal progesterone) was offered if cervical length
129 was ≤ 25mm or based on clinician and patient preference in case of a large change in cervical
130 length between measurements. None of the low-risk women required an intervention but
131 71.6% (78/109) of the high-risk women did (Table 1). Initiation of this treatment was always
132 after the VMB sampling visits.
133
134 *VMB laboratory assessments*
135 DNA was extracted from one swab per woman per visit using lysozyme lysis and bead-
136 beating procedures combined with the Qiagen DNeasy Blood and Tissue kit (Qiagen,

Manchester, UK) (Appendix A).[13] The V3-V4 region of the 16S rRNA gene was amplified and sequenced on an Illumina HiSeq 2500 instrument (Illumina, San Diego, CA) run in rapid mode, 2x300bp using a 250PE and 50E kit. The panbacterial 16S rRNA gene copy concentration per sample was determined using the BactQuant quantitative PCR (qPCR) assay.[14] Molecular data processing steps are described in Appendix A.

We only used the first set of valid sequencing results for each participant in the analyses because of the short time period between the two study visits, with one exception. In the subset of participants with valid sequencing results for both visits (88 high-risk and 129 low-risk women) we used data from both visits for the VMB stability analyses.

Creation and selection of VMB variables

Existing approaches

We systematically reviewed the published literature to identify VMB variables that were associated with PTB in at least one study (Appendix A, Table A1). These included: richness and Simpson diversity (1-D) as continuous variables; two categorical VMB composition variables with each participant assigned to one category (the community state types (CSTs) described by Ravel et al[15] and *Lactobacillus* groups (dominant, intermediate, and deplete) based on *Lactobacillus* relative abundance);[16] stability groups, slightly modified from Romero *et al*;[17] and the presence/absence of specific taxa of interest (listed in Table A1).

New approaches

The final three sets of VMB variables are based on our previous work and were applied to PTB research for the first time.[18–21] The first is a categorical VMB composition variable that

we named 'VMB type'. These VMB types are similar to Ravel's CSTs[15] but with an increased emphasis on the relative abundance of pathobionts (Appendix A). We defined pathobionts as bacteria that are considered more pathogenic than BV-anaerobes and that often co-occur with lactobacilli instead of BV-anaerobes, such as streptococci. Secondly, each non-minority amplicon sequence variant (ASV; equivalent to a taxon) in each sample was allocated to one of four 'bacterial groups' based on the published literature (Appendix B): lactobacilli; BV-anaerobes; pathobionts; and a rest group of 'other bacteria'.[20] Within each sample, read counts of ASVs belonging to the same bacterial group were summed. This resulted in four continuous relative abundance variables (one for each bacterial group) per sample, which sum to 1.0 for each sample. Finally, we converted these bacterial group relative abundances into estimated concentrations (again four continuous variables, one for each bacterial group) making use of the BactQuant results (Appendix A). The BactQuant data were also used to determine estimated concentrations of taxa of interest (listed in Table A1) and the total bacterial load in each sample.

Statistical analysis

Our primary comparisons were between high-risk women who gave birth $\geq 37^{+0}$ weeks gestation and high-risk women who had a recurrent early sPTB or PPROM $<34^{+0}$ weeks. Low-risk women who delivered at $\geq 39^{+0}$ weeks gestation without PPROM were used as the normal reference group because we had never before characterised VMB compositions in pregnant women residing in Liverpool. Characteristics between these three groups were compared by student's t-test for age, Mann-Whitney U test for other continuous variables, and Fisher's exact test for binary and categorical variables. The various VMB variables described above were compared between the two high-risk groups using unadjusted and adjusted logistic regression, with adjustments for body mass index (BMI; as a quadratic term

due to the bimodal association with PTB), history of cervical surgery (none, single LLETZ, or multiple LLETZ or knife cone biopsy), and smoking at the time of enrolment (yes/no). Quartiles of the expected distributions of total vaginal bacterial load and total estimated *Lactobacillus* concentration were generated using data from the low-risk reference group, and high-risk participants were allocated to one of these quartiles. These newly created categorical variables were also compared between the two high-risk groups of interest by logistic regression. Finally, in an effort to differentiate between the effects of total vaginal bacterial load and the types of bacteria that make up this load, women were stratified by VMB type, and the logistic regression analyses were repeated for each stratum.

Results

We recruited 137 high-risk women and 223 low-risk women. Of 131 high-risk women with VMB data that passed quality control, 22 had a recurrent early sPTB/PPROM and 87 had term births ≥37^{+0} weeks without PPROM (Figure 1); the remaining women were excluded because they delivered between 34^{+0} and 37^{+0} weeks or had a birth initiated by a caregiver. Of 214 low-risk participants with VMB data that passed quality control, 145 gave birth at ≥39^{+0} weeks without PPROM and were used as the normal reference group.

The participant characteristics of the three groups were similar (Tables 1 and A2), except for those that are known risk factors for sPTB/PPROM: a higher proportion of high-risk women with a recurrence, compared to those who delivered at term, had two or more previous early sPTB/PPROM events (31.8% vs. 9.2%) and multiple previous LLETZ or knife cone biopsies (13.6% vs. 1.1%). The median gestational age when the first valid VMB sample was taken was slightly later in the low-risk group (16^{+5} weeks) than in the high-risk group (16^{+3}

210 weeks).

211

212 *VMB compositions and vaginal bacterial loads in high- and low-risk pregnant women*

213 The VMBs of over 70% of pregnant women, high- and low-risk, were dominated by

214 lactobacilli (defined as ≥75% relative abundance; Tables 2 and A3, heatmap in Figure A1). *L.*

215 *iners* domination was most common overall (VMB type Li; 33.0% of high-risk and 22.1% of

216 low-risk pregnant women), followed by *L. crispatus* domination (VMB type Lcr; 21.1% and

217 26.9%), and either domination by other lactobacilli (most commonly *L. gasseri* or *L. jensenii*)

218 or ≥50% bifidobacteria (VMB type Lo+BL; 20.1% and 25.5%). In addition, 12.4% of high-risk

219 and 12.8% of low-risk women had mild anaerobic dysbiosis (VMB type LA; defined as 25%-

220 75% lactobacilli with the remainder BV-anaerobes), and 12.8% of high-risk and 13.1% of low-

221 risk women had severe anaerobic dysbiosis (VMB type BV; ≥75% BV-anaerobes). Pathobionts

222 and 'other bacteria' were rarely present, and if present, only in low relative abundance or

223 estimated concentration (Tables 2 and A3).

224

225 The VMB of the majority of participants with samples available from both visits between

226 15^{+0} and 23^{+0} weeks gestation was stable (Table 2). About two thirds of the high-risk (65.9%)

227 and low-risk women (64.3%) were dominated by the same *Lactobacillus* species at both

228 visits, and 14.8% of high-risk and 22.5% of low-risk women continued to have some degree

229 of anaerobic dysbiosis. The remaining women switched *Lactobacillus* species (n=5) or shifted

230 from lactobacilli-domination to anaerobic dysbiosis (n=12) or vice-versa (n=17).

231

232 The median vaginal bacterial load in the low-risk reference cohort was 7.68 $\log_{10}$ cells/µl

233 (interquartile range (IQR) 6.80-8.35) and in the high-risk cohort 8.07 (IQR 7.28-8.92) $\log_{10}$

234 cells/µl (Table 2). High-risk women with mild or severe anaerobic dysbiosis (pooled together)
235 had higher median vaginal bacterial loads than women with domination by *L. iners, L.*
236 *crispatus*, or other lactobacilli/bifidobacteria (8.54, 7.79, 7.63, 7.53 $\log_{10}$ cells/µl,
237 respectively; p-values are <0.05 for all comparisons between dysbiosis and the other groups,
238 but are not significant for the comparisons between *L. iners*-domination and the other two
239 lactobacilli groups; Table A4).
240
241 *Associations with early sPTB/PPROM recurrence – existing approaches*
242 In the high-risk women, we did not identify any associations between recurrence and VMB
243 richness, Simpson (1-D) diversity, CSTs,[15] categorical *Lactobacillus* groups,[16] categorical
244 stability groups[17] (Table 2), and presence or relative abundance of common *Lactobacillus*
245 species (Table A3). *Ureaplasma* species and BVAB TM7-H1 were present in higher
246 proportions of high-risk women with a recurrence than in high-risk women with a term birth,
247 and *Dialister* species vice versa, but their relative abundances when present were very low
248 (Table A3). In contrast, *Bifidobacterium breve* was present in a lower proportion of women
249 with a recurrence, but its relative abundance when present was also low. The proportions of
250 low-risk women with these taxa present were similar to those in high-risk women without a
251 recurrence (Table A3, Figure A2A).
252
253 *Associations with early sPTB/PPROM recurrence – new approaches*
254 We did not identify any associations between recurrence and VMB type or estimated BV-
255 anaerobes or pathobionts concentrations (Table 2). However, high-risk women with a
256 recurrence had a higher overall vaginal bacterial load (8.64 vs. 7.89 $\log_{10}$ cells/µl, adjusted
257 odds ratio (aOR) 1.90, 95% confidence interval (CI) 1.01-3.56, p=0.047) and a higher total

258 estimated *Lactobacillus* concentration (8.59 vs. 7.48 $\log_{10}$ cells/µl, aOR 2.35, 95% CI 1.20-
259 4.61, p=0.013) than high-risk women without a recurrence (Table 2). The aOR was 1.63 (95%
260 CI 0.93-2.86, p=0.093) for each increase in total vaginal bacterial load quartile, and 2.61 (95%
261 CI 1.27-5.33, p=<0.001) for each estimated *Lactobacillus* concentration quartile (Table 2).
262
263 Table 3 shows estimated concentrations of total vaginal bacteria, the four bacterial groups,
264 *L. cripatus*, *L. iners*, and other lactobacilli in high-risk women with and without a recurrence
265 after stratification by VMB type. A trend towards higher recurrence risk with higher median
266 bacterial loads was seen for each VMB type after stratification but was only significant in
267 women with *L. iners*-domination: the median vaginal bacterial loads were 9.03 (IQR 8.02-
268 9.12) $\log_{10}$ cells/µl in women with a recurrence compared to 7.79 (IQR 6.93-8.65) $\log_{10}$
269 cells/µl in women with a term birth (aOR 3.44, 95%CI 1.06-11.15, p=0.011). The association
270 was no longer significant after adjustment for confounders (aOR 2.38, 95%CI 0.71-7.95,
271 p=0.160). Within each of the other VMB type strata, the median estimated concentrations in
272 high-risk women who had a term birth were similar to those in low-risk women, and non-
273 significantly lower than those in high-risk women who had a recurrence.
274
275 The majority of high-risk participants with an early sPTB/PPROM recurrence (17/22; 77.3%),
276 and about half of the high-risk participants without a recurrence (44/87; 50.6%), had at least
277 one VMB characteristic that was positively associated with recurrence. The following
278 positively associated VMB characteristics overlapped considerably in both high-risk groups
279 (Venn diagrams in Figure A2): highest quartile of total bacterial load, highest quartile of
280 estimated total *Lactobacillus* concentration, highest quartile of estimated *L. iners*
281 concentration, and presence of BVAB TM7-H1. The only positively associated VMB

282 characteristic that did not completely overlap with the highest quartile of bacterial and
283 estimated total *Lactobacillus* concentration was the presence of *Ureaplasma* species (this
284 was the case in high-risk pregnancies with and without a recurrence).
285

286 ## Discussion

287 *Main Findings*

288 This exploratory study suggests that high vaginal bacterial load in the second trimester is
289 associated with early sPTB/PPROM recurrence in high-risk pregnant women, irrespective of
290 the bacterial composition. Our statistical power was limited: only 22 women had an early
291 sPTB/PPROM recurrence. Therefore, not all associations in line with this conclusion reached
292 statistical significance, especially after reducing statistical power further by stratification.
293 However, we believe that this new hypothesis warrants further research for various reasons.
294 First, the recurrence risk increased for each increase in bacterial and *Lactobacillus*
295 concentration quartile (on a $\log_{10}$ scale). Second, the findings were consistent for each VMB
296 type after stratification: higher recurrence in high-risk women with higher median bacterial
297 loads for all VMB types (reaching significance for *L. iners*-domination), and similar median
298 bacterial loads in high-risk and low-risk women who had a term birth for all VMB types.
299 Third, estimated concentrations reached higher levels for anaerobic dysbiosis and *L. iners*-
300 domination than for domination by other lactobacilli or bifidobacteria. Our findings
301 therefore do not contradict earlier findings by other groups that indicated increased PTB risk
302 for anaerobic dysbiosis and high *L. iners* relative abundance.[8,22–24] In fact, the inconsistent
303 findings related to *L. iners* in the literature might be due to the lack of quantification in most
304 studies.
305

306 *Interpretation in light of other evidence*

307 To our knowledge, the potential association between the vaginal bacterial load and PTB has
308 only been assessed in two previous studies. Freitas et al[7] also showed a higher 16S gene
309 concentration in the second trimester (approximately 16 weeks gestation) in women who
310 had an sPTB <37 weeks compared to women with a term birth ≥37 weeks (8.1 vs. 7.8 $\log_{10}$
311 copies/swab, respectively). Conversely, Elovitz et al[8] reported a similar 16S gene
312 concentration in women who had a sPTB <37 weeks compared to women with a term birth
313 ≥37 weeks in samples collected before 24 weeks gestation (approximately 8 $\log_{10}$
314 copies/swab in each group), and a lower concentration (approximately 7.5 vs 7.9 $\log_{10}$
315 copies/swab) in samples collected at 28 weeks gestation. High vaginal bacterial load may
316 only be associated with PTB when present prior to 24 weeks gestation, but additional
317 quantitative studies are needed.

318

319 Female genital tract inflammation is considered one potential cause of PTB.[25] BV (a clinical
320 condition caused by anaerobic dysbiosis) is associated with vaginal inflammation, and is a
321 very common condition in non-pregnant women.[26] Pregnant women are protected from BV
322 due to the high oestrogen levels during pregnancy.[27,28] In our study, the prevalence of
323 anaerobic dysbiosis, and the relative abundances and estimated concentrations of BV-
324 associated bacteria and pathobionts, were indeed low. This likely explains why we did not
325 identify any statistically significant associations between the presence or extent of anaerobic
326 dysbiosis and early sPTB/PPROM recurrence. However, the non-significant trends
327 consistently suggested increased risk. Another common inflammatory vaginal condition is
328 vulvovaginal candidiasis, for which pregnant women are at increased risk.[29] In our study,
329 only one woman (low-risk) reported having been treated for vulvovaginal candidiasis prior to

330 study enrolment. Future studies should systematically assess vaginal yeasts concentrations
331 in addition to bacterial concentrations in all participants, and determine their potential
332 contributions to early sPTB/PPROM recurrence.
333
334 In our study, *L. iners* was able to achieve higher bacterial loads than the other lactobacilli,
335 and was associated with early sPTB/PPROM recurrence. *L. iners* has a genome of just 1.3
336 Mbp,[30] strikingly smaller than the other common vaginal *Lactobacillus* species, and seems
337 fully adapted to the vaginal niche.[31] This may explain why it is able to grow to such high
338 densities in the vagina. *Lactobacillus* species are considered to be non-inflammatory when
339 present in the vagina,[32] but they might cause inflammation when they travel through the
340 cervical mucus plug into the upper genital tract. Ascension to the upper genital tract may be
341 increased when vaginal density is high. Studies have shown that *L. iners* with high expression
342 of clustered regularly inter-spaced short palindromic repeat (CRISPR) genes are present in
343 women with BV but not in women without BV.[33] CRISPR genes are the primary bacterial
344 defence against bacteriophages. The *L. iners* strains associated with high bacterial loads and
345 recurrent sPTB/PPROM in our study may differ from strains associated with lower bacterial
346 loads. This requires further research.
347
348 Two uncommon taxa (present in fewer than 20% of the women in the low risk group) were
349 non-significantly associated with early sPTB/PPROM recurrence in our study and were also
350 associated with PTB in other studies: *Ureaplasma*[7,24] and BVAB TM7-H1[6]. In women with
351 early sPTB/PPROM recurrence, the presence of *Ureaplasma* species in the vagina did not
352 completely overlap with having high bacterial loads and lactobacilli concentrations.
353 *Ureaplasma* species may therefore play a role in PTB that is independent from overall

vaginal bacterial composition and load. Our data confirms that *Bifidobacterium breve* may have a protective effect against early sPTB/PPROM recurrence.[16]

Strengths and Limitations

Our study had some limitations, most notably limited statistical power and insufficient data on vaginal yeasts, but also had strengths. Pregnancy outcomes were very carefully selected and assessed, many VMB composition variables were created covering different aspects of VMB composition, and most of these VMB variables were (semi)quantified. The latter turned out to be crucial, as our most important findings concern bacterial concentrations. The quantification method that we used has recently been validated for non-minority taxa in two different studies.[7,8] We had initially planned to analyse early sPTB and early PPROM as two separate outcomes because the aetiological pathways may differ,[34] and to exclude women who used PTB prevention treatments,[35] but the limited statistical power did not permit this. A final limitation is the lack of VMB data in the third trimester and/or closer to the birth.

Conclusion

Vaginal bacterial load in the second trimester, irrespective of the bacterial composition, was associated with early sPTB/PPROM recurrence. Domination by lactobacilli other than *L. iners* may protect women from developing high bacterial loads. These findings should be confirmed in larger, longitudinal studies that incorporate quantification of vaginal bacteria and yeasts. If they are confirmed, interventions that maintain a non-*iners* lactobacilli/ bifidobacteria-dominated VMB may protect women from inflammation-associated PTB.

Acknowledgements

We would like to thank all participants for their enthusiastic involvement in the study, in particular members of the Harris-Wellbeing Patient and Public Engagement group. We would also like to thank the Liverpool Women's Hospital for hosting the research, Tracy Ricketts at the Liverpool Women's Hospital for administrative support, University of Liverpool Centre for Genomic Research staff for 16S sequencing, and Jacques Ravel and Mike Humphrys at the University of Maryland School of Medicine, Institute for Genome Sciences, for the 16S qPCR assays.

Disclosure of interests

All authors have completed the ICMJE uniform disclosure forms and declare: LG, AC, AS, AA, BM-M, AA, ZA and JW received a grant from Wellbeing of Women charity to establish the Harris Wellbeing Research Centre that funded the submitted work. MW, aCG, CB, DR, JI and BP received no support from any organisation for the submitted work. All authors declare no financial relationships with any organizations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Contribution to authorship

JW, AS, DR, BM-M, AC, AA and ZA conceived the study, wrote the protocol and obtained funding. AC, JI, BP and LG contributed to the protocol and recruited participants. AD, JW, aCG, and CB developed the DNA extraction, sequencing, and bioinformatics methods. LG, AC and AS extracted the clinical data. LG performed the laboratory analysis. JW developed the VMB analysis methods, and LG and MV performed the data processing. LG performed the data analysis and wrote the initial draft, and JW, ZA and MV contributed to data interpretation and revised the paper. All authors approved the final manuscript.

Details of Ethics Approval

The study was approved by North West Research Ethics Committee - Liverpool Central, reference 11/NW/0720 on 4 November 2011. The collection of vaginal fluid for this component of the study was approved in amendment 4 on 19 October 2015.

Funding

The prospective cohort study was funded Wellbeing of Women as part of a charitable donation from Lord and Lady Harris to establish the Harris-Wellbeing PTB Research Centre, University of Liverpool. This covered administrative costs, laboratory analysis, salary for AC, study support costs for AC and LG. No additional funding was used. Wellbeing of Women were not involved in the conduct of the research or writing the paper.

References

1. Ferrero DM, Larson J, Jacobsson B, Di Renzo GC, Norman JE, Martin JN, et al. Cross-Country Individual Participant Analysis of 4.1 Million Singleton Births in 5 Countries with Very High Human Development Index Confirms Known Associations but Provides No Biologic Explanation for 2/3 of All Preterm Births. Luo Z-C, editor. PLoS One [Internet]. 2016 Sep 13 [cited 2019 Mar 11];11(9):e0162506. Available from: https://dx.plos.org/10.1371/journal.pone.0162506
2. Zhang G, Srivastava A, Bacelis J, Juodakis J, Jacobsson B, Muglia LJ. Genetic studies of gestational duration and preterm birth. Best Pract Res Clin Obstet Gynaecol [Internet]. 2018 Oct [cited 2019 Mar 11];52:33–47. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30007778
3. Hwang J, Vinturache AE, Jacobsson B, Gyamfi-Bannerman C, Mysorekar IU. Maternal microbiome – A pathway to preterm birth. Semin Fetal Neonatal Med. 2016;21(2):94–9.
4. Romero R, Hassan SS, Gajer P, Tarca AL, Fadrosh DW, Bieda J, et al. The vaginal microbiota of pregnant women who subsequently have spontaneous preterm labor and delivery and those with a normal delivery at term. Microbiome [Internet]. 2014 Dec 27 [cited 2019 Mar 19];2(1):18. Available from: https://microbiomejournal.biomedcentral.com/articles/10.1186/2049-2618-2-18
5. Peelen MJ, Luef BM, Lamont RF, de Milliano I, Jensen JS, Limpens J, et al. The influence of the vaginal microbiota on preterm birth: A systematic review and recommendations for a minimum dataset for future research. Placenta [Internet]. 2019;79(March):30–9. Available from: https://doi.org/10.1016/j.placenta.2019.03.011
6. Fettweis JM, Serrano MG, Brooks JP, Edwards DJ, Girerd PH, Parikh HI, et al. The vaginal microbiome and preterm birth. Nat Med. 2019;25(6):1012–21.
7. Freitas AC, Bocking A, Hill JE, Money DM. Increased richness and diversity of the vaginal microbiota and spontaneous preterm birth. Microbiome. 2018;6(1):1–15.
8. Elovitz MA, Gajer P, Riis V, Brown AG, Humphrys MS, Holm JB, et al. Cervicovaginal microbiota and local immune response modulate the risk of spontaneous preterm delivery. Nat Commun [Internet]. 2019;10(1):1–8. Available from: http://dx.doi.org/10.1038/s41467-019-09285-9
9. Jian C, Luukkonen P, Yki-Jarvinen H, Salonen A, Korpela K. Quantitative PCR provides a simple and accessible method for quantitative microbiome profiling. bioRxiv [Internet]. 2018;478685. Available from: https://www.biorxiv.org/content/early/2018/11/27/478685
10. Tettamanti Boshier FA, Srinivasan S, Lopez A, Hoffman NG, Proll S, Fredricks DN, et al. Complementing 16S rRNA Gene Amplicon Sequencing with Total Bacterial Load To Infer Absolute Species Concentrations in the Vaginal Microbiome. mSystems. 2020;5(2):1–14.
11. Van't Hooft J, Duffy JMN, Daly M, Williamson PR, Meher S, Thom E, et al. A core outcome set for evaluation of interventions to prevent preterm birth. Obstet Gynecol. 2016;127(1):49–58.
12. National Institute for Health and Care Excellence. Antenatal care for uncomplicated pregnancies. Clin Guidel [Internet]. 2008;(March):1–55. Available from: https://www.nice.org.uk/guidance/cg62/resources/antenatal-care-for-uncomplicated-pregnancies-975564597445
13. Gill C, Van De Wijgert JHHM, Blow F, Darby AC. Evaluation of lysis methods for the extraction of bacterial DNA for analysis of the vaginal microbiota. PLoS One. 2016;11(9):1–16.
14. Liu CM, Aziz M, Kachur S, Hsueh PR, Huang YT, Keim P, et al. BactQuant: an enhanced broad-coverage bacterial quantitative real-time PCR assay. BMC Microbiol. 2012;12.
15. Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SSK, McCulle SL, et al. Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci U S A. 2011;108(SUPPL. 1):4680–7.
16. Brown RG, Al-Memar M, Marchesi JR, Lee YS, Smith A, Chan D, et al. Establishment of vaginal microbiota composition in early pregnancy and its association with subsequent preterm prelabor rupture of the fetal membranes. Transl Res [Internet]. 2019;1–14. Available from: https://doi.org/10.1016/j.trsl.2018.12.005

17. Romero R, Hassan SS, Gajer P, Tarca AL, Fadrosh DW, Nikita L, et al. The composition and stability of the vaginal microbiota of normal pregnant women is different from that of non-pregnant women. Microbiome. 2014;2(1):1–19.
18. van de Wijgert JHHM, Verwijs MC. Lactobacilli-containing vaginal probiotics to cure or prevent bacterial or fungal vaginal dysbiosis: a systematic review and recommendations for future trial designs. BJOG An Int J Obstet Gynaecol. 2019;1–13.
19. Verwijs MC, Agaba SK, Darby AC, van de Wijgert JHHM. Impact of Oral Metronidazole Treatment on the Vaginal Microbiota and Correlates of Treatment Failure. Am J Obstet Gynecol [Internet]. 2019;(August). Available from: https://linkinghub.elsevier.com/retrieve/pii/S0002937819310075
20. van de Wijgert JHHM, Verwijs MC, Gill AC, Borgdorff H, van der Veer C, Mayaud P. Pathobionts in the Vaginal Microbiota: Individual Participant Data Meta-Analysis of Three Sequencing Studies. Front Cell Infect Microbiol. 2020;10(April).
21. van de Wijgert JHHM, Verwijs MC, Agaba SK, Bronowski C, Mwambarangwe L, Uwineza M, et al. Intermittent Lactobacilli-containing Vaginal Probiotic or Metronidazole Use to Prevent Bacterial Vaginosis Recurrence: A Pilot Study Incorporating Microscopy and Sequencing. Sci Rep [Internet]. 2020;10(1):1–15. Available from: http://dx.doi.org/10.1038/s41598-020-60671-6
22. Kindinger LM, Bennett PR, Lee YS, Marchesi JR, Smith A, Cacciatore S, et al. The interaction between vaginal microbiota, cervical length, and vaginal progesterone treatment for preterm birth risk. Microbiome [Internet]. 2017;5(1):1–14. Available from: http://dx.doi.org/10.1186/s40168-016-0223-9
23. Callahan BJ, DiGiulio DB, Goltsman DSA, Sun CL, Costello EK, Jeganathan P, et al. Replication and refinement of a vaginal microbial signature of preterm birth in two racially distinct cohorts of US women. Proc Natl Acad Sci U S A [Internet]. 2017;114(37):9966–71. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28847941%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5604014
24. DiGiulio DB, Callahan BJ, McMurdie PJ, Costello EK, Lyell DJ, Robaczewska A, et al. Temporal and spatial variation of the human microbiota during pregnancy. Proc Natl Acad Sci U S A. 2015;112(35):11060–5.
25. Romero R, Espinoza J, Gonçalves LF, Kusanovic JP, Friel L, Hassan S. The Role of Inflammation and Infection in Preterm Birth. Semin Reprod Med. 2007;25(01):21–39.
26. Kenyon C, Colebunders R, Crucitti T. The global epidemiology of bacterial vaginosis: A systematic review. Am J Obstet Gynecol [Internet]. 2013;209(6):505–23. Available from: http://dx.doi.org/10.1016/j.ajog.2013.05.006
27. Cruickshank R, Sharman A. The Biology of the Vagina in the Human Subject. BJOG An Int J Obstet Gynaecol [Internet]. 1934 Apr 1;41(2):208–26. Available from: https://doi.org/10.1111/j.1471-0528.1934.tb08759.x
28. Hay P. Bacterial Vaginosis as a Mixed Infection. In: Brogden K, Guthmiller J, editors. Polymicrobial Diseases [Internet]. Washington (DC); 2002. Available from: https://www.ncbi.nlm.nih.gov/books/NBK2495
29. Fidel PL, Sobel JD. Immunopathogenesis of recurrent vulvovaginal candidiasis. Clin Microbiol Rev. 1996;9(3):335–48.
30. Macklaim JM, Gloor GB, Anukam KC, Cribby S, Reid G. At the crossroads of vaginal health and disease, the genome sequence of Lactobacillus iners AB-1. Proc Natl Acad Sci U S A. 2011;108(SUPPL. 1):4688–95.
31. Van Der Veer C, Hertzberger RY, Bruisten SM, Tytgat HLP, Swanenburg J, De Kat Angelino-Bart A, et al. Comparative genomics of human Lactobacillus crispatus isolates reveals genes for glycosylation and glycogen degradation: Implications for in vivo dominance of the vaginal microbiota. Microbiome. 2019;7(1):1–14.
32. Könönen E. Anaerobic Cocci and Anaerobic Gram-Positive Nonsporulating Bacilli. In: Bennett

J, Dolin R, Blaser M, editors. Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases. 8th ed. Chicago: Elsevier Inc.; 2015. p. 2781-2786.e2.

33. Macklaim JM, Fernandes AD, Di Bella JM, Hammond JA, Reid G, Gloor GB. Comparative meta-RNA-seq of the vaginal microbiota and differential expression by Lactobacillus iners in health and dysbiosis. Microbiome. 2013;1(1):1–11.

34. Capece A, Vasieva O, Meher S, Alfirevic Z, Alfirevic A. Pathway analysis of genetic factors associated with spontaneous preterm birth and pre-labor preterm rupture of membranes. PLoS One. 2014;9(9).

35. Peelen MJ, Luef BM, Lamont RF, de Milliano I, Jensen JS, Limpens J, et al. The influence of the vaginal microbiota on preterm birth: A systematic review and recommendations for a minimum dataset for future research. Placenta [Internet]. 2019;79(December 2018):30–9. Available from: https://doi.org/10.1016/j.placenta.2019.03.011

Tables, Figures, and Supplementary Materials Caption List

Table 1: Participant characteristics by pregnancy risk group and outcome

Table 2: Associations between VMB variables and recurrent early sPTB/PPROM in high-risk women

Table 3: Associations between estimated median concentrations of vaginal bacteria and early recurrent sPTB/PPROM stratified by VMB type

Figure 1: Participant flow

Appendix A: Supplementary methods, figures, and tables

Appendix B: Assignment of amplicon sequence variants to bacterial groups

Table 1: Participant characteristics by pregnancy risk group and outcome

		Low-risk pregnancy (LR)		High-risk pregnancy (HR)				P* value LR vs HR	P* value HR term vs HR preterm
		Term birth n=145		Term birth n=87		Early sPTB or PPROM n=22			
Age (years)	mean (SD)	31.1	(4.5)	30.0	(4.5)	31.4	(5.4)	0.145	0.211
BMI (kg/m2)¶	median (IQR)	24.5	(22-29)	25	(22-28)	27.5	(22-35)	0.619	0.139
Current smoker¶	number (%)	15	(10.6)	15	(17.4)	5	(23.8)	0.096	0.537
Ethnicity (%)	White	139	(95.9)	81	(93.1)	22	(100.0)	0.540	0.732
	Black	3	(2.1)	4	(4.6)	0	0.0		
	Asian	0	0.0	0	0.0	0	0.0		
	Other	2	(1.4)	0	0.0	0	0.0		
	Not recorded	1	(0.1)	2	(2.3)	0	0.0		
Parity (%)	0	0	0.00	9	(10.34)	1	(4.55)	0.000	0.017
	1	102	(70.34)	43	(49.43)	10	(45.45)		
	2	36	(24.83)	27	(31.03)	4	(18.18)		
	≥ 3	7	(4.83)	8	(9.20)	7	(31.82)		
Number of previous PPROM or sPTB 16+0-33+6 weeks (%)	1			79	(90.8)	15	(68.2)	na	0.012
	≥ 2			8	(9.2)	7	(31.8)		
Previous cervical surgery	Nil significant	131	(90.3)	79	(90.8)	17	(77.3)	na	0.033
	Single LLETZ	14	(9.7)	7	(8.0)	2	(9.1)		
	Multiple LLETZ or knife cone			1	(1.1)	3	(13.6)		
Gestational age at first valid VMB sample (weeks)	median and IQR	16+5	(16+2-17+1)	16+2	(16+0-16+6)	16+3	(15+6-17+3)	0.000	0.794
Cervical length¶ (mm) (median and IQR)	16 weeks	41	(36-47.1)	35.5	(31-42)	36	(26.5-43.5)	0.000	0.578
	20 weeks	41	(37-45)	35	(30-40)	30	(26-37)	0.000	0.097
qfFN¶ (ng/ml) (median and IQR)	16 weeks	7	(5-16)	7	(5-22)	9	(6-23.5)	0.766	0.410
	20 weeks	7	(5-12)	7	(5-14)	7	(6-12)	0.274	0.914
Preterm birth prevention treatment used after study visit	None	139	(100.0)	61	(70.1)	17	(77.3)	na	0.097
	Cervical cerclage			4	(4.6)	2	(9.1)		
	Progesterone			2	(2.3)	1	(4.5)		
	Arabin pessary			20	(23.0)	2	(9.1)		
Antimicrobial in pregnancy prior to sampling (%)	None	115	(79.3)	65	(74.7)	17	(77.3)	0.038	0.895
	Metronidazole or clindamycin	0	0.0	4	(4.6)	1	(4.5)		
	Other antibiotic	21	(14.5)	9	(10.3)	3	(13.6)		
	Clotrimazole	1	(0.7)	0	0.0	0	0.0		
	Not recorded/ unsure	8	(5.5)	9	(10.3)	1	(4.5)		
Gestational age at PPROM (weeks)	median and range					31+1	18+0-33+6	na	na
Gestational age at birth (weeks)	median and range	40+1	(39+0-41+6)	38+6	37+0-41+5	31+5	18+0-35+5		
Birthweight¶ (g)	mean (SD)	3594	(439)	3234	(489)	1778	(673)		

Abbreviations: BMI, body mass index; HR, high-risk pregnancy; LR, low-risk pregnancy; na, not applicable; PPROM, preterm prelabour rupture of membranes; qfFN, quantitative foetal fibronectin level in cervicovaginal fluid; sPTB, spontaneous preterm birth.

* P values by Student's t-test for age, Mann-Whitney U test for BMI, cervical length, qfFN and gestational age at sampling, and Fisher's exact test for remainder of variables.

¶Missing values: BMI 1 each for LR and HR term; smoking 3 for LR, and 1 each for HR term and early preterm; birthweight 2 each for HR term and HR preterm. Cervical length at 16 weeks: HR term n=80, HR preterm n=20, LR n=136. Cervical length at 20 weeks: HR term n=79, HR preterm n=14, LR n=135. qfFN at 16 weeks: HR term n=77, HR preterm n=20, LR n=142. qfFN at 20 weeks: HR term n=75, HR preterm n=19, LR n=136.

Parity was defined as number of previous pregnancies with live births or stillbirths ≥24^{+0} weeks gestation. Eligibility for the HR cohort included ≥1 sPTB or PPROM at 16^{+0} to 33^{+6} weeks gestation. Therefore, all HR women with parity=0 had at least one previous late miscarriage. Antimicrobial usage was only assessed in the pregnancy prior to sampling. Clotrimazole is available without medical review as an 'over the counter' medication in the UK, and so usage after the study visit was not possible to assess within our study design.

Table 2: Associations between VMB variables and recurrent early sPTB/PPROM in high-risk women

Basis for assessment			Low-risk pregnancy (LR)		High-risk pregnancy (HR)						p value* difference between HR ≥ 37 vs. <34 weeks	Adjusted¶ logistic regression model of HR ≥ 37 vs <34 weeks (n=106)		
			Birth ≥39 weeks n=145		All high-risk pregnancies n=109		Birth ≥37 weeks n=87		sPTB or PPROM <34 weeks n=22					
			Median (IQR)	n (%)	Median (IQR)	n %	Median (IQR)	n %	Median (IQR)	n %		OR	95% CI	P
Relative abundance data (16S rRNA sequencing)	Global assessments	Richness (n)	7 (4-14)		7 (4-15)		7 (4-15)		8 (5-14)		0.35	0.99	(0.93-1.06)	0.785
		Simpson diversity (1-D)	0.38 (0.07-0.53)		0.30 (0.06-0.53)		0.39 (0.08-0.54)		0.19 (0.05-0.45)		0.60	0.29	(0.04-2.18)	0.232
	VMB type⁋	Lcr		39 (26.9)		23 (21.1)		17 (19.5)		6 (27.3)	0.696	Comparator		
		Li		32 (22.1)		36 (33.0)		28 (32.2)		8 (36.4)		0.59	(0.15-2.34)	0.455
		Lo+BL		37 (25.5)		22 (20.2)		17 (19.5)		5 (22.7)		0.81	(0.17-3.79)	0.793
		LA		18 (12.4)		14 (12.8)		13 (14.9)		1 (4.5)		0.16	(0.01-1.98)	0.154
		BV		19 (13.1)		14 (12.8)		12 (13.8)		2 (9.1)		0.48	(0.07-3.24)	0.451
	Community State Type	I *(L. crispatus)*		43 (29.7)		23 (21.1)		17 (19.5)		6 (27.3)	0.848	Comparator		
		II *(L. gasseri)*		16 (11.0)		11 (10.1)		9 (10.3)		2 (9.1)		0.38	(0.04-3.95)	0.421
		III *(L. iners)*		31 (21.4)		33 (30.3)		26 (29.9)		7 (31.8)		0.55	(0.13-2.24)	0.402
		IV (mixed non Lactos)		35 (24.1)		29 (26.6)		25 (28.7)		4 (18.2)		0.43	(0.09-2.01)	0.287
		V *(L. jensenii)*		20 (13.8)		13 (11.9)		10 (11.5)		3 (13.6)		0.94	(0.17-5.12)	0.944
	Group according to Lactobacillus relative abundance	*Lactobacillus* dominance (>75%)		105 (72.4)		80 (73.4)		61 (70.1)		19 (86.4)	0.277	Comparator		
		Intermediate (25-74.9%)		14 (9.7)		6 (5.5)		5 (5.7)		1 (4.5)		0.91	(0.09-9.36)	0.940
		Lactobacillus deplete (<25%)		26 (17.9)		23 (21.1)		21 (24.1)		2 (9.1)		0.29	(0.05-1.53)	0.140
	Stability: change in VMB between 16 and 20 weeks gestation⁋	Same Lactobacilli	n=129	83 (64.3)	n=88	58 (65.9)	n=71	46 (64.8)	n=17	12 (70.6)	0.579	Comparator		
		Different Lactobacilli		3 (2.3)		2 (2.3)		1 (1.4)		1 (5.9)		4.40	(0.19-104)	0.359
		Remain non-LD		29 (22.5)		13 (14.8)		12 (16.9)		1 (5.9)		0.42	(0.04-4.11)	0.457
		Non-LD to LD		7 (5.4)		10 (11.4)		8 (11.3)		2 (11.8)		0.96	(0.16-5.90)	0.965
		LD to non-LD		7 (5.4)		5 (5.7)		4 (5.6)		1 (5.9)		0.41	(0.02-7.21)	0.543
Relative abundance combined with BactQuant, producing estimated concentration in log10 cells/µl	Global assessment	Bacterial load	7.68 (6.80-8.35)	145 (100.0)	8.07 (7.28-8.92)	109 (100.0)	7.89 (7.22-8.80)	87 (100.0)	8.64 (7.47-9.08)	22 (100.0)	**0.035**	1.90	(1.01-3.56)	**0.047**
	Bacterial group	Total Lactobacillus	7.31 (6.42-8.01)	143 (98.6)	7.38 (6.76-8.17)	107 (98.2)	7.48 (6.87-8.45)	85 (97.7)	8.59 (7.47-8.87)	22 (100.0)	**0.005**	2.35	(1.20-4.61)	**0.013**
		Total BV associated	5.41 (1.91-7.03)	109 (75.2)	5.49 (0-7.65)	81 (74.3)	5.20 (0-7.65)	62 (71.3)	6.65 (5.09-8.36)	19 (86.4)	0.102	1.15	(0.96-1.38)	0.126
		Total pathobionts	0.00 (0-0)	34 (23.4)	0.00 (0-0)	24 (22.0)	0.00 (0-0)	20 (23.0)	0.00 (0-0)	4 (18.2)	0.596	0.93	(0.712-1.22)	0.596
		Total other bacteria	0.00 (0-4.24)	50 (34.5)	0.00 (0-4.93)	45 (41.3)	0.00 (0-4.84)	35 (40.2)	0.00 (0-5.12)	10 (45.5)	0.671	1.04	(0.86-1.25)	0.698
	Type of Lactobacillus	*L. crispatus*	5.88 (4.41-6.83)	125 (86.2)	6.22 (4.44-6.95)	89 (81.7)	6.17 (4.08-6.91)	71 (81.6)	6.32 (5.79-7.76)	18 (81.8)	0.319	1.02	(0.83-1.25)	0.843
		L. iners	4.65 (0-7.14)	82 (56.6)	5.47 (0-7.33)	76 (69.7)	6.55 (0-7.98)	63 (72.4)	7.42 (0-8.81)	13 (59.1)	0.684	0.97	(0.84-1.12)	0.644
		Other Lactobacilli	5.49 (3.77-6.91)	125 (86.2)	5.94 (4.24-7.07)	99 (90.8)	5.81 (4.21-7.09)	78 (89.7)	5.95 (5.34-7.61)	21 (95.5)	0.172	1.14	(0.88-1.47)	0.311
	Allocation of high risk participants to quartiles of low risk		Concentration											
	Quartile of total bacterial load (n, %)	1	<6.80	37 (25.5)		12 (11.0)		11 (12.6)		1 (4.5)	0.093	1.63	(0.93-2.86)	0.087
		2	≥6.80-<7.68	36 (24.8)		34 (31.2)		28 (32.2)		6 (27.3)				
		3	≥7.68-<8.35	35 (24.1)		18 (16.5)		17 (19.5)		1 (4.5)				
		4	≥8.35	37 (25.5)		45 (41.3)		31 (35.6)		14 (63.6)				
	Quartile of total Lactobacillus (n, %)	1	<6.42	36 (24.8)		13 (11.9)		12 (13.8)		1 (4.5)	**0.018**	2.61	(1.27-5.33)	**0.009**
		2	≥6.42- >7.31	36 (24.8)		25 (22.9)		24 (27.6)		1 (4.5)				
		3	≥7.31- >8.01	36 (24.8)		29 (26.6)		23 (26.4)		6 (27.3)				
		4	≥8.01	37 (25.5)		42 (38.5)		28 (32.2)		14 (63.6)				

Abbreviations: HR, high-risk pregnancy; LR, low-risk pregnancy; PPROM, preterm prelabour rupture of membranes; sPTB, spontaneous preterm birth; VMB, vaginal microbiota.

*P values by Mann-Whitney U test for continuous variables and Fisher's exact test for remainder. ¶Logistic regression adjusted for body mass index as quadratic term, history of cervical surgery (none/single LLETZ/multiple LLETZ and/or knife cone biopsy) and smoking (yes/no at time of visit). Adjusted analyses excluded two HR women with missing smoking data, one with missing BMI data, and variable numbers because of collinearity within the regression. ⁋VMB type definitions: Li=*L. iners*-dominated (≥75% lactobacilli with *L. iners* the most common); Lcr=*L. crispatus*-dominated (≥75% lactobacilli with *L. crispatus* the most common); Lo+BL=other lactobacilli- or *Bifidobacterium* dominated (Lo+BL; either≥75% lactobacilli with *L. jensenii* or *L. gasseri* the most common, or ≥50% *Bifidobacterium*); lactobacilli and anaerobes (LA; 25%- 75% lactobacilli); and BV=mixture of BV-anaerobes (<25% lactobacilli). ⁋VMB changes between 16 and 20 weeks were only assessed in the subset of participants with valid VMB assessments at both time points: LR n=129; n=71 HR term n=71, and HR preterm n=17.

Table 3: Associations between estimated median concentrations of vaginal bacteria and early recurrent sPTB/PPROM stratified by VMB type

VMB type¶			Low risk pregnancy (LR)		High risk pregnancy (HR)						Unadjusted logistic regression model of HR ≥ 37 vs <34 weeks (n=109)			Adjusted* logistic regression model of HR ≥ 37 vs <34 weeks (n=106)		
			Birth ≥ 39 weeks n=145		All high-risk pregnancies n=109		Birth ≥ 37 weeks n=87		sPTB or PPROM <34 weeks (early preterm) n=22							
			Median (IQR)		Median (IQR)		Median (IQR)		Median (IQR)		OR	95% CI	P	OR	95% CI	P
LCr		n	39		23		17		6		23			18		
	Bacterial load		7.75	(6.44-8.46)	7.63	(7.14-8.63)	7.55	(7.14-8.32)	8.20	(7.42-9.06)	1.67	(0.57-4.89)	0.349	1.35	(0.41-4.40)	0.623
	Bacterial group	Total Lactobacillus	7.73	(6.42-8.46)	7.63	(7.14-8.59)	7.55	(7.14-8.25)	8.20	(7.41-8.86)	1.64	(0.55-4.85)	0.375	1.32	(0.40-4.30)	0.649
		Total BV associated	4.74	(0-6.29)	4.62	(0-6.07)	4.29	(0-5.36)	5.24	(0-6.71)	1.09	(0.79-1.50)	0.601	0.87	(0.55-1.36)	0.535
		Total pathobionts	0.00	(0-0)	0.00	(0-0)	0.00	(0-0)	0.00	(0-0)	0.93	(0.49-1.75)	0.825	1.07	(0.48-2.39)	0.866
		Total other bacteria	0.00	(0-3.19)	0.00	(0-0)	0.00	(0-0)	0.00	(0-0)	1.01	(0.59-1.71)	0.980	0.97	(0.56-1.68)	0.904
	Type of Lactobacillus	*L. crispatus*	7.72	(6.41-8.45)	7.63	(7.12-8.37)	7.55	(7.12-8.22)	8.19	(7.41-8.80)	1.66	(0.55-5.02)	0.365	1.31	(0.40-4.27)	0.658
		L. iners	0.00	(0-4.23)	0.00	(0-4.81)	0.00	(0-5.32)	0.00	(0-0)	0.71	(0.41-1.21)	0.204	0.84	(0.49-1.42)	0.514
		Other Lactobacilli	5.09	(3.39-5.94)	5.34	(3.02-7.16)	4.64	(3.02-6.14)	6.20	(5.34-7.39)	1.17	(0.77-1.76)	0.457	1.03	(0.93-1.80)	0.930
Li		n	32		36		28		8		36			34		
	Bacterial load		7.64	(6.94-8.25)	7.79	(6.93-8.65)	7.79	(6.93-8.65)	9.03	(8.02-9.12)	3.44	(1.06-11.15)	**0.040**	2.38	(0.71-7.95)	0.160
	Bacterial group	Total Lactobacillus	7.61	(6.90-8.25)	7.96	(7.23-8.83)	7.79	(6.93-8.60)	8.91	(8.02-9.10)	3.40	(1.04-11.11)	**0.042**	2.37	(0.70-8.04)	0.165
		Total BV associated	4.55	(0-6.37)	5.16	(0-6.60)	4.75	(0-5.79)	6.77	(5.33-7.73)	1.45	(0.98-2.15)	0.061	1.57	(0.95-2.59)	0.079
		Total pathobionts	0.00	(0-0)	0.00	(0-0)	0.00	(0-0)	0.00	(0-2.31)	1.21	(0.84-1.74)	0.307	1.15	(0.69-1.91)	0.590
		Total other bacteria	0.00	(0-3.61)	0.00	(0-4.26)	0.00	(0-3.38)	4.07	(0-5.69)	1.39	(1.00-1.94)	0.050	1.43	(0.95-2.16)	0.085
	Type of Lactobacillus	*L. crispatus*	5.89	(3.75-6.60)	6.19	(4.45-6.84)	6.02	(3.83-6.78)	6.38	(6.12-7.79)	1.26	(0.82-1.94)	0.295	1.09	(0.69-1.72)	0.695
		L. iners	7.52	(6.90-8.12)	7.93	(7.13-8.79)	7.73	(6.93-8.59)	8.82	(8.02-9.07)	3.35	(1.03-10.86)	**0.044**	2.33	(0.70-7.76)	0.169
		Other Lactobacilli	4.97	0-6.75)	5.59	(4.08-6.59)	5.40	(4.41-7.30)	5.81	(4.08-6.59)	1.20	(0.78-1.85)	0.398	1.06	(0.69-1.63)	0.782
Lo or BL		n	37		22		17		5		22			19		
	Bacterial load		7.12	(6.32-7.69)	7.53	(7.24-8.64)	7.48	(7.21-8.73)	7.64	(7.46-8.61)	1.36	(0.48-3.86)	0.563	3.47	(0.61-19.7)	0.161
	Bacterial group	Total Lactobacillus	7.12	(6.08-7.69)	7.45	(7.10-8.61)	7.36	(6.94-7.89)	7.61	(7.42-8.61)	1.68	(0.54-5.29)	0.371	5.96	(0.65-54.5)	0.114
		Total BV associated	4.64	(0-5.59)	4.29	(0-6.19)	0.00	(0-4.94)	6.36	(6.31-6.58)	1.84	(0.97-3.50)	0.063	model failed		
		Total pathobionts	0.00	(0-2.98)	0.00	(0-4.11)	0.00	(0-4.11)	0.00	(0-0)	0.87	(0.53-1.41)	0.560	0.66	(0.23-1.87)	0.430
		Total other bacteria	0.00	(0-5.01)	1.13	(0-5.93)	0.00	(0-5.95)	5.09	(0-5.12)	1.05	(0.76-1.44)	0.769	1.21	(0.77-1.91)	0.410
	Type of Lactobacillus	*L. crispatus*	5.17	(4.00-5.88)	5.55	(4.44-6.54)	5.52	(5.26-6.54)	5.59	(0-6.16)	0.85	(0.59-1.21)	0.362	0.83	(0.53-1.29)	0.410
		L. iners	0.00	(0-6.41)	0.00	(0-7.09)	4.42	(0-7.09)	0.00	(0-0)	0.84	(0.61-1.16)	0.297	0.87	(0.58-1.30)	0.497
		Other Lactobacilli	6.99	(5.97-7.51)	7.46	(6.84-8.51)	7.24	(6.93-7.84)	7.61	(7.42-8.30)	1.87	(0.58-6.05)	0.293	8.97	(0.66-121)	0.099
LA or BV		n	37		28		25		3		28			27		
	Bacterial load		7.93	(7.44-8.52)	8.54	(8.11-9.22)	8.31	(8.07-9.16)	9.15	(9.11-9.28)	9.46	(0.28-318)	0.210	12.42	(0.31-501)	0.182
	Bacterial group	Total Lactobacillus	6.98	(6.30-7.45)	7.38	(6.63-8.37)	7.25	(6.56-8.19)	8.26	(8.15-8.86)	4.57	(0.54-38.5)	0.162	5.49	(0.17-63.1)	0.172
		Total BV associated	7.93	(7.19-8.51)	8.41	(8.07-9.19)	8.18	(8.07-9.15)	9.05	(8.83-9.24)	6.58	(0.33-133)	0.219	9.93	(0.42-237)	0.156
		Total pathobionts	0.00	(0-0)	0.00	(0-3.92)	0.00	(0-4.49)	0.00	(0-0)	model failed			model failed		
		Total other bacteria	0.00	(0-4.40)	4.57	(0-6.07)	4.59	(0-5.96)	0.00	(0-6.30)	0.83	(0.54-1.28)	0.402	0.70	(0.30-1.62)	0.407
	Type of Lactobacillus	*L. crispatus*	4.98	(0-5.93)	5.21	(0-6.41)	4.88	(0-6.60)	5.79	(0-5.90)	1.04	(0.69-1.55)	0.859	0.84	(0.48-1.46)	0.537
		L. iners	5.65	(0-7.11)	6.68	(4.80-8.21)	6.56	(4.74-7.98)	8.26	(8.15-8.86)	4.74	(0.49-45.6)	0.179	9.82	(0.35-274)	0.179
		Other Lactobacilli	4.79	(2.37-5.84)	5.32	(4.12-6.04)	5.34	(4.01-6.06)	5.20	(5.10-5.44)	1.09	(0.64-1.85)	0.756	1.03	(0.55-1.93)	0.916

Abbreviations: 95% CI, 95% confidence interval; HR, high-risk pregnancy; IQR, interquartile range; LR, low-risk pregnancy; OR, odds ratio; PPROM, preterm prelabour rupture of membranes; sPTB, spontaneous preterm birth; VMB, vaginal microbiota. *Logistic regression adjusted for body mass index as quadratic term, history of cervical surgery (none/single LLETZ/multiple LLETZ and/or knife cone biopsy) and smoking (yes/no at time of visit). Adjusted analyses excluded two HR women with missing smoking data, one with missing BMI data, and variable numbers because of collinearity within the regression. Participants were included in these analyses even if a taxa was not present in that sample (with concentration=0). ¶VMB type definitions: Li=*L. iners*-dominated (≥75% lactobacilli with *L. iners* the most common); Lcr=*L. crispatus*-dominated (≥75% lactobacilli with *L. crispatus* the most common); Lo+BL=other lactobacilli- or *Bifidobacterium* dominated (Lo+BL; either≥75% lactobacilli with *L. jensenii* or *L. gasseri* the most common, or ≥50% *Bifidobacterium*); lactobacilli and anaerobes (LA; 25%- 75% lactobacilli); and BV=mixture of BV-anaerobes (<25% lactobacilli).

High-risk pregnancy

(Previous sPTB or PPROM <34 weeks)

Low-risk pregnancy

(Parous women with all previous births at ter

N=137

N=223

Numbers	Exclusions	Numbers
NA	Early term delivery 37^{+0}-38^{+6} (low risk only)	65
14	sPTB/PPROM 34^{+0} -37^{+0}	2
NA	sPTB/PPROM <34^{+0} (low risk only)	2
8	Caregiver initiated preterm birth	0
6	VMB data failed quality control	9

N=109

N=145

gh-Risk Early Preterm

n=22

sPTB or PPROM

<34^{+0} weeks

High-Risk Term

n=87

Birth ≥37^{+0} weeks

Low-Risk Term

n=145

Birth ≥39^{+0} weeks