Group B Streptococcus and the Vaginal Microbiome among Pregnant Women: a Systematic Review

Sungju Lim¹, Shilpa Rajagopal¹, Ye Ryn Jeong¹, Dumebi Nzegwu¹, and Michelle Wright²

¹The University of Texas at Austin
²University of Texas at Austin

November 3, 2020

Abstract

Background: Current analysis methods of the vaginal microbiome often fail to identify and/or report known pathogens such as S. agalatiae (GBS), resulting in an incomplete clinical understanding of bacterial composition during pregnancy. Objective: To determine the frequency of GBS reporting in vaginal microbiome literature pertaining to pregnancy and examine methodological bias that contributes to differences in species and genus-level microbiome reporting. Search strategy: PubMed, CINAHL, and Web of Science databases were used to identify papers for review based on the search terms “vaginal microbiome”, “pregnancy”, and “16S rRNA sequencing”. Selection criteria: Forty-five research articles reported employing a metagenomic approach for vaginal microbiome analysis during pregnancy that explicitly reported taxonomic composition and were included in this review. Data collection and analysis: Articles were evaluated for methods of DNA extraction and sequencing, 16S region, taxonomy classification database, number of participants or vaginal specimens, and pregnancy trimester. Main results: Less than 30% of articles reported the presence of GBS. No significant differences in methodology were identified between articles that reported versus did not report GBS. However, large variability across the types of research methods used for vaginal microbiome analysis and species-level bacterial community reporting were observed. Conclusion: Considerable differences in study design and data formatting methods may contribute to underrepresentation of GBS, and other known pathogens, in existing vaginal microbiome literature. Providing detailed methods, analysis code and raw data may improve reproducibility and ability to compare microbial communities across studies. Keywords: vaginal microbiome, pregnancy, 16S ribosomal RNA, whole genome sequencing, Group B Streptococcus
Objective: To determine the frequency of GBS reporting in vaginal microbiome literature pertaining to pregnancy and examine methodological bias that contributes to differences in species and genus-level microbiome reporting.

Search strategy: PubMed, CINAHL, and Web of Science databases were used to identify papers for review based on the search terms “vaginal microbiome”, “pregnancy”, and “16S rRNA sequencing”.

Selection criteria: Forty-five research articles reported employing a metagenomic approach for vaginal microbiome analysis during pregnancy that explicitly reported taxonomic composition and were included in this review.

Data collection and analysis: Articles were evaluated for methods of DNA extraction and sequencing, 16S region, taxonomy classification database, number of participants or vaginal specimens, and pregnancy trimester.

Main results: Less than 30% of articles reported the presence of GBS. No significant differences in methodology were identified between articles that reported versus did not report GBS. However, large variability across the types of research methods used for vaginal microbiome analysis and species-level bacterial community reporting were observed.

Conclusion: Considerable differences in study design and data formatting methods may contribute to underrepresentation of GBS, and other known pathogens, in existing vaginal microbiome literature. Providing detailed methods, analysis code and raw data may improve reproducibility and ability to compare microbial communities across studies.

Keywords vaginal microbiome, pregnancy, 16S ribosomal RNA, whole genome sequencing, Group B Streptococcus

Tweetable abstract: Heterogeneity in methods and taxonomy reporting may underestimate known pathogens in the vaginal microbiome.

Introduction

Rationale

Variations within the vaginal microbiome have been associated with higher risk preterm birth and other pregnancy complications such as chorioamnionitis. Microbes associated with these complications such as Streptococcus agalactiae and Gardnerella vaginalis, are frequently present in polymicrobial as well as Lactobacillus predominant vagial microbiomes. In light of these considerations, a growing body of research has focused on characterizing the vaginal microbiome using metagenomic approaches. Among the most common methodologies are those that rely on universal targets, 16S ribosomal RNA and cpn60 genes, allowing for more detailed taxonomic classification and profiling of the vaginal microbiome during pregnancy than culture based or targeted methods. However, data from amplicon-based metagenomic characterizations suggest that such methods provide an incomplete understanding of the species-level diversity present in the vaginal microbiome. For example, there is a notable lack of discussion of S. agalactiae, commonly referred to as Group B Streptococcus (GBS), a known pathogen associated with poor maternal, fetal, and neonatal outcomes in amplicon-based metagenomic vaginal microbiome studies during pregnancy.

Currently, in the United States, screening for GBS occurs at 36 to 38 gestational weeks, with those who test positive receiving antibiotic treatment during labor to help prevent early-onset GBS. Previous data indicate that approximately one in four women is colonized by GBS (Centers for Disease Control and Prevention). Yet, apart from this routine clinical screening near the time of birth, there are limited metagenomic studies that identify or evaluate GBS during pregnancy. While screening for GBS prior to birth has reduced early onset neonatal sepsis caused by GBS in the U.S., the transient nature of maternal GBS colonization is not well understood. Improved understanding of the dynamics of maternal GBS colonization during pregnancy can provide additional insights for preventing GBS complications affecting maternal and fetal health throughout pregnancy and birth (e.g., miscarriage and stillbirth).
Objectives
To our knowledge, there is no systematic review addressing lack of representation of known pathogens, such as GBS, associated with pregnancy complications in the existing vaginal microbiome literature. Therefore, this systematic review was conducted to: 1) determine how often GBS is reported in studies of the vaginal microbiome during pregnancy, and 2) analyze methods used to identify potential methodological bias for GBS reporting or underrepresentation in vaginal microbiome studies during pregnancy.

Methods
Protocol and Registration
This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic reviews and Meta-analyses (PRISMA) guidelines.7

Eligibility Criteria
Original studies that investigated the vaginal microbiome during human pregnancy, were written in English, and published within the past 10 years were included. Only studies with pregnant women as study participants and reported microbiome taxonomy were included. Review papers, editorials, commentaries, and methodological papers including procedure/protocols were excluded. Research studies that used a targeted approach (i.e.: only on specific microbial taxa, such as Lactobacillus species, were analyzed via microbe specific PCR), were also excluded.

Information Sources
Three electronic databases PubMed, CINAHL, and Web of Science were used for data extraction (date last searched: June 30, 2020).

Search
We searched these databases using two groups of search terms: (“vaginal microbiome” or “vaginal microbiota” or “pregnancy”) and (“16S ribosomal RNA” or “16S rRNA” or “whole genome sequencing”). For example, in PubMed, (((vaginal microbiome) OR (vaginal microbiota))) AND pregnancy) AND ((16s ribosomal rRNA) OR (16s rRNA) OR (whole genome sequencing) OR (metagenomic)) were entered in the search window with the display option ‘best match’. Limitations of publication date as ‘10 years’, species as ‘humans’, and language as ‘English’ were used (Appendix S1).

Study Selection
After combining articles from the three databases and removing duplicates, five reviewers (SL, SR, YJ, DN, MLW) independently screened all results with titles and abstracts. Disagreements were resolved by discussion among all five reviewers to reach consensus. After the screening, eligibility was checked with full text by five reviewers (SL, SR, YJ, DN, MLW), who each reviewed one fifth of the total articles and cross-reviewed another one fifth. Disagreements were resolved by discussion to reach consensus.

Data Collection Process
Data was extracted from the included studies by using a collaborative Google Sheets document. Each of five reviewers (SL, SR, YJ, DN, MLW) extracted one fifth of the included studies and cross-checked the accuracy of extracted data from another one fifth. Ambiguity was resolved through discussion among all reviewers.

Data Items
The initial data extraction tool contained the items as follows: study title, study purpose, whether GBS was reported, level of analysis - genus/species (in the case of GBS not reported), sequencing methods, 16S rRNA primer region (for studies using 16S rRNA sequencing), sequencing platform, taxonomy classification database, DNA collection kit, DNA extraction kit, number of pregnant women, number of vaginal specimens,
age of pregnant women, pregnancy period in which the specimen collected, ethnicity/region of the partici-
pants/study design. In studies that reported taxonomy to the species level and GBS was not detected, we
also assessed the articles to determine if *Streptococcus* at the genus level was reported.

Risk of Bias in Individual Studies

No previous studies have exclusively examined the methodology of vaginal microbiome studies related to
detection of known common opportunistic pathogens, like GBS. One of the main purposes of this review
was to evaluate if methodological differences/biases in vaginal microbiome studies contribute to reporting or
underrepresentation of GBS.

Summary Measures

Our primary outcome was to identify the number of studies that reported the presence of GBS within the
vaginal microbiome. Due to methodological heterogeneity, the differences across analytic and sequencing
methodologies between studies that reported versus did not report GBS were compared narratively.

Synthesis of Results

After data extraction, the included studies were divided into two groups: GBS reported and GBS not
reported. Independent t-tests were conducted to compare the number of pregnant women and the number
of vaginal specimens between GBS reported and not reported groups. The commonalities and differences
among 16S rRNA primer regions used were compared. We then qualitatively compared the commonalities
and differences of the other methodological categories between two groups, since methods were too varied to
compare statistically.

Results

Study Selection

Our search strategy initially yielded 278 manuscripts (Figure 1). After removing 68 duplicates, 210 article
titles and abstracts were screened, and 159 studies were excluded because they did not meet inclusion criteria.
Full text was reviewed for the remaining 52 manuscripts, and 6 studies additional studies excluded for failing
to meet inclusion criteria. A total of 45 studies were included and reviewed in this study.

Study Characteristics and Results of Individual Studies

The characteristics of each study that reported (n = 13) or did not report GBS (n = 32) are described in
the Supplementary Table S1 and S2 respectively.

Synthesis of Results

Among the 45 studies reviewed, less than 30% of manuscripts reported the presence of GBS within the
results, figures, or microbial taxonomy tables. To determine the frequency in which GBS was reported, and
to compare analytic methods across studies, we analyzed: time of sample collection, number of research
participants and samples, DNA extraction method, type of sequencing completed, and database used to
assign taxonomy (Table 1).

Participant and sample characteristics

There were no differences in the sample sizes (mean comparison of the number of specimens: t = 0.372, df
= 43, p = 0.712; number of women: t = 0.862, df = 43, p = 0.394). However, there was large variability in
sample size and number of specimens collected in studies that reported and did not report GBS. Additionally,
four manuscripts that did not report GBS within the taxonomy analyzed, or overall metagenomic results,
did report positive culture-based GBS clinical screening results for women within the study.8–11

DNA extraction kits, Sequencing method, and Taxonomy classification

A variety of approaches were employed for DNA extraction; custom developed protocols, along with two
standard extraction kits (Table 1) were used in studies reporting GBS within the vaginal microbiome.
However, these kits and custom approaches were also used in studies that did not report GBS. In terms of sequencing methods, all studies reporting GBS used 16S rRNA sequencing, as well as most studies that did not report the presence of GBS. Additionally, there was no notable difference found in the 16S rRNA hypervariable region used across either set of studies. Similarly, multiple taxonomy databases were used for studies that did and did not report the presence of GBS.

Risk of Bias Across Studies

The included studies were not originally conducted with a specific focus on GBS. However, most studies were evaluating the composition of the microbiome across pregnancy, or related to birth outcomes that GBS can influence, such as preterm birth (Supplementary Tables S1 and S2). There is notable bias against GBS reporting when only genus level data are reported, as well as when studies only speciate lactobacillus. For example, seven studies that did not report GBS only assigned taxonomy to the genus level, 9 only assigned species level assignments for Lactobacillus spp., three only assigned species to Lactobacillus spp., and Gardnerella, and three studies assigned taxonomy to species level, that were included in analysis, but did not report species level results (Table S1). Given that species other than lactobacillus can negatively impact vaginal health and birth outcomes, these methods of analysis and reporting contribute to positive bias towards Lactobacillus reporting in the vaginal microbiome literature and negative bias towards other relevant species. Further, four studies only reported the top 15-21 most abundant species or taxa (Table S1). Since some vaginal pathogens present in low abundance can cause pregnancy complications the underrepresentation of these pathogens, such as GBS, may indicate bias against reporting of microbes that are not the most abundant taxonomic groups.

Discussion

Main Findings

In this review, we explored how often GBS was reported in studies of the vaginal microbiome during pregnancy. Methods related to GBS reporting were compared to identify potential methodological bias. Among 45 studies, 13 (28.9%) reported GBS. However, due to the heterogeneity of methods across study approaches, we were not able to detect any systematic differences that appeared to enhance or hinder GBS reporting.

Strengths and Limitations

Strengths of this review include that both outcomes and methods were used to analyze the vaginal microbiome composition. We identified studies from multiple databases and assessed differences across studies at various decision points to analyze metagenomic data.

The heterogeneity across studies in methods and data reporting limited the ability to pool and re-analyze the data. This was further complicated by the fact that some studies reported most microbial taxa at solely the genus level, whereas other studies only included species-level data for specific genus like Lactobacillus. Furthermore, the format and availability of data associated with manuscripts also varied greatly (e.g., raw data versus including only taxonomy tables used in the analysis). In cases where the full microbial datasets were not available, it was difficult to determine whether the absence of GBS reporting was due to lack of detection, exclusion from analysis because of low abundance, or some other factor. These limitations highlight the importance of making study analysis code and raw data freely available upon publication in order to encourage reproducibility and comparison.

Interpretation

Methodologic differences and sources of bias

There are known sources of bias in metagenomic studies associated with analysis methods. Brooks and colleagues determined that microbiome community composition can be biased at various steps including DNA extraction, PCR amplification, sequencing, and taxonomic classification. The researchers reported that more than 85% of the microbiome community results were biased by small variations in methods (less than 5%). Specifically, it has been determined that variations in DNA extraction and amplification can
result in decreased representation of *Streptococcus* species. If women with GBS colonization only harbor small amounts of the bacteria, it is possible that bias introduced by these steps could have contributed to the limited representation of GBS in vaginal microbiome data.

Such underrepresentation of known pathogens can significantly limit the clinical scope of vaginal microbiome studies. For example, many studies only reported results related to community state types or analyzed the most abundant species. Consequently, if thresholds for inclusion are based on sample abundance (e.g., 5% of sample or 5,000 reads), GBS data may have been excluded if it did not meet the threshold.

Furthermore, selection of 16S regions for amplification and associated primers can influence genus and species level resolution. For example, it is generally more difficult to distinguish *Lactobacillus* species within the V4 region, but may be easier to do so in V1-V3. Conversely, *G. vaginalis* is more difficult to detect and differentiate within the V1-V2 region because its sequence is highly variable prior to the V1 region, making primer selection more complex. These biases can be partially mitigated by recognizing the characteristics of 16S rRNA sequencing and thoughtfully selecting primers. Unfortunately, studies that evaluate the influence of selected 16S regions on GBS representation have not been published yet. From our findings, the 13 studies that detected GBS used 16S regions between V1-V4 or V6. However, the 32 studies that did not report GBS also used the same regions. Thus, assessment of bias by 16S region is important when designing and evaluating future vaginal microbiome studies.

Given limitations in exploring the true proportions within the microbial community, studies have increasingly employed whole genome sequencing (WGS), which allows for higher resolution up to the strain level. In our review, two studies employed WGS but did not report GBS. Both papers initially selected some of the most frequent species-level taxa, which are frequently *Lactobacillus* and *Gardnerella* species, to continue further analysis. This approach may have constrained the strength of WGS, causing researchers to miss the rarer, yet clinically important genus, species, and strains.

**Limited assessment and lack of focus on low abundance organisms**

The above results indicate that reporting strategies, rather than methodology, may contribute to the underreporting of less abundant pathogens in metagenomic vaginal microbiome studies during pregnancy. As the majority of studies in our review explored the compositional variation of vaginal microbiome during pregnancy, underreporting can be a serious issue because known pathogens may not be analyzed as potential contributors to pregnancy complications, which can lead to limited clinical treatment.

We observed differences in threshold determinants for analysis or reporting. Among the 32 studies that did not report GBS, 24 analyzed the microbiome at the species level, but only reported the most prevalent genus (i.e., *Lactobacillus* species). This may lead to a misclassification error, in which only the most abundant bacteria or species is associated with the outcome of interest. Additionally, 20 studies reported *Streptococcus* at the genus level, but did not conduct further species-level analysis. It is possible that studies that did not specify *Streptococcus* had GBS present within their population, a result which could be re-evaluated if raw sequencing data is made available.

Moreover, four studies that did not report GBS in their results or microbiome data identified GBS-positive women in their study based on culture-based clinical testing. This suggests that there may be a benefit to using both sequence-based and culture-based methods to verify presence or absence of low abundance microbes within vaginal microbiome communities. Previous studies have shown discordance in culture-based GBS colonization results between antenatal screening and culture on admission for delivery. In a study by McCoy and colleagues, that rate was about 40%. Alternatively, Hussain et al. reported a discordance rate of 11%. While the transient nature of GBS colonization is known, factors that contribute to loss or gain of maternal GBS colonization remain uncertain.

Evaluation of culture-based GBS screening determined that GBS transmission to neonates is more likely when maternal colonization is heavy, as determined by the number of colonies grown at the time of screening. However, it is unclear how well clinical culture-based microbial assessments correlate with the detection of
clinically relevant pathogens in sequencing-based vaginal microbiome studies. Furthermore, much of the research investigating risk of neonatal GBS infection related to maternal bacterial load was completed in the 1980s and has not been evaluated relative to sequencing-based methods. Future studies that concurrently evaluate the accuracy of GBS reporting within microbial communities via metagenomic sequencing approaches and culture-based methods are needed.

Conclusion

Only a small proportion of vaginal microbiome studies reported the presence of GBS during pregnancy. No systematic differences in factors previously known to introduce bias in microbiome studies were associated with the lack of GBS reporting. However, there was considerable heterogeneity in research methods employed across studies. Consideration of less abundant, but clinically meaningful microbes in vaginal microbiome studies may improve our understanding of how the vaginal microbiota influences pregnancy outcomes. GBS was reported in studies across all trimesters of pregnancy using metagenomic methods, although reporting of culture based GBS does not always correlate with metagenomic GBS reporting. Therefore, studies evaluating the presence of GBS may need to adopt both confirmatory testing for GBS via culture and parallel comparison with results obtained from metagenomic/meta-taxonomic approaches.

Disclosure of Interests:

None declared. Completed disclosure of interests form available to view online as supporting information.

Contribution to Authorship

SL, SR and MLW designed and planned the study. SL, SR and MLW performed the search. SL, SR, YJ, DN, and MLW independently screened and selected eligible studies, and completed data extraction, verifying each other’s work. SL, SR, YJ and MLW contributed to the interpretation of the data and the writing of the article. All authors reviewed and approved the final version of the manuscript and accepted responsibility for the paper as published.

Details of Ethics Approval

Not applicable.

Funding: MLW is supported by the National Institute of Nursing Research (K01NR017903). The content is solely the responsibility of the author and does not necessarily represent the official views of the NIH.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Details of the included studies – GBS reported studies (n = 13)Table S2. Details of the included studies – GBS not reported studies (n = 32)Appendix S1. Database search strategy.

References


Table 1. Comparison of GBS Reported and GBS Not Reported Studies

<table>
<thead>
<tr>
<th>Trimesters</th>
<th>GBS Reported (n=13)</th>
<th>GBS Not Reported (n=32)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st trimester (1)</td>
<td>1st trimester (5)</td>
</tr>
<tr>
<td></td>
<td>2nd trimester (0)</td>
<td>2nd trimester (6)</td>
</tr>
<tr>
<td></td>
<td>3rd trimester (3)</td>
<td>3rd trimester (6)</td>
</tr>
<tr>
<td></td>
<td>All trimesters (5)</td>
<td>All trimesters (12)</td>
</tr>
<tr>
<td></td>
<td>2nd and 3rd (2)</td>
<td>1st and 2nd (2)</td>
</tr>
<tr>
<td></td>
<td>Unspecified (2)</td>
<td>2nd and 3rd (5)</td>
</tr>
</tbody>
</table>

| Number of pregnant women | Mean: 111.154 (SD 121.817) | Mean: 82.594 (SD 91.331) |
|                         | Range [8 - 450], Total # 1445 | Range [6 - 461], Total # 2643 |

| Number of vaginal specimens | (11/13 studies; 2 unspecified) | (32/32 studies) |
|                            | Mean: 184 | Mean: 214 |
|                            | Standard deviation: 155.793 | Standard deviation: 418.706 |
|                            | Range: [8 - 450] | Range: [4 - 2179] |
|                            | Total #: around 2086 | Total #: around 6848 |

| DNA extraction kit          | QIAamp DNA Mini Kit (6) | QIAamp (6) |
|                            | MoBio PowerSoil Kit (3) | MoBio (13) |
|                            | Custom (4)             | DNeasy (2) |
|                            | Custom (5)             | Others (5) |

| Sequencing method          | 16S rRNA Sequencing (13) | 16S rRNA Sequencing (29) |
|                            | Shotgun sequencing (2)  | Cpn60 (2) |

| 16S rRNA region            | V1-V2 (4), V1-V3 (3), V3-V4 (1), V4 (3), V6 (1), full (1) | V1-V2 (2), V1-V3 (8), V2 (1), V3 (8), full (5) |

Hosted file

Figure.pdf available at https://authorea.com/users/372543/articles/490538-group-b-streptococcus-and-the-vaginal-microbiome-among-pregnant-women-a-systematic-review