Author’s response to reviews

Title: Relationship between Alterations of Urinary Microbiota and Cultured Negative Lower Urinary Tract Symptoms in Female Type 2 Diabetes Patients

Authors:

Jiawei Chen (miwaicjw@163.com)
Jie Zhao (zhaojie-0412@163.com)
Ying Cao (nfcy123@126.com)
Guihao Zhang (jyzhanggh@163.com)
Yang Chen (cyhm28@sina.com)
Jialei Zhong (15625171604@163.com)
Weina Huang (1287823961@qq.com)
Jiarong Zeng (15521266641@163.com)
Peng Wu (doctorwupeng@gmail.com)

Version: 1 Date: 24 Nov 2018

Author’s response to reviews:

Replies to Cindy L. Amundsen (Reviewer 1):

Abstract:
1. Page 2, lines 26-27: The aim of your abstract is too vague - you should be more specific about how you are studying alterations of urinary microbiota and why. Moreover, I don't think you can really say that your aim is to "explore the pathogenesis" of culture-negative LUTS because this is a cross sectional study. You cannot infer causality (pathogenesis) based on a cross-sectional study. Thus, your stated aim is the wrong one for this study.

Answer: Changes have been made as suggested.

(Abstract section, line 24-28, page 2): Background: Lower urinary tract symptoms (LUTS) is the most common complication of diabetes. However, the underlying pathogenesis of cultured negative LUTS (cn-LUTS) in diabetic patients has not been well understood. Numerous evidence indicates that urinary dysbiosis is related to urologic disorders. We aim to study alterations of the urinary microbiota of cn-
LUTS in type 2 diabetes (T2D) patients.

2. Page 2, line 29: It seems that you used clean catch urine specimens for this study (although you should clarify whether that is true or not). Why did you not use catheterized urine samples? There are numerous urinary microbiome studies indicating that clean catch urine samples are more reflective of the peri-urethral and vaginal areas rather than the bladder.

Answer: Changes have been made as suggested. These issues have been supplemented in the discussion section and method section.

(Methods section, line 119-121, page 6): Mid-stream urine (50ml) was collected with the labial separation which was supervised by the author (Weina Huang) and 20ml was left for standard cultivation to exclude UTI.

(Discussion section, line 369-377, page 17-18): Owing to the urine samples may be contaminated with microbiota surrounding the urethral orifice, mid-stream urine was collected by the clean catch method with labial separation supervised by the author (Weina Huang). Catheter-derived urine samples were suggested as an alternative method, but the method was considered being invasive and was not ethically workable in patients who do not have a clinical indication for it. Comparison of results from Siddiqui et al.'s urinary microbiota study on female mid-stream urine with the consequences of suprapubic aspirate by Wolfe et al. shown that the major findings were the same [12]. It was a strong indication that results from mid-stream urinary microbiota were reliable.

3. Your abstract should have a separate "methods" subsection. You have included some of your methods in the results section of the abstract and it's better to separate these into sub-sections since that's what the readers will expect.

Answer: Changes have been made as suggested. We initially submitted our manuscript to BMC Microbiology. Following the journal style of BMC Microbiology, there is no "methods" subsection in the abstract. However, editor of BMC microbiology considered our manuscript was not suitable for the journal and invited us to use the transfer service to transfer our manuscript to BMC Urology. Thus, the abstract was lack of "methods" subsection.

(Discussion section, line 29, page 2): Methods:

4. Page 2, lines 28-44: You haven't presented any raw numbers or proportions here, but only p values. Please present the raw numbers of your results. This also applies to the full results section of the paper. For example, don't just say that 2 values are "significantly different." You should present the raw numbers of each and state (for example) that one value was statistically decreased (or increased) compared to the other.

Answer: Changes have been made as suggested. As for the beta diversity, it reflects the differences in microbial communities between groups, thus, we can't use “statistically decreased (or increased)” to describe differences in beta diversity. Revisions of the results section are shown in Question 16.

(Abstract section, line 35-41, page 2): However, statistically decreased richness (ACE index and Chao 1 index, 85.52(13.75, 204.84) vs. 129.82(63.89, 280.30) and 83.86(11.00, 210.77) vs. 125.19(62.00, 251.77), P=0.005; Observed Species, 76(10, 175) vs. 98(54, 234), P=0.011) and decreased species diversity (Shannon index, 1.37(0.04, 3.48) vs. 2.09(0.98, 3.43), P = 0.033;
Simpson index, 0.46 (0.06, 0.99) vs. 0.23 (0.07, 0.64), P = 0.029) were shown in moderate-to-severe LUTS group and high Hemoglobin A1c group, respectively.

5. Page 3, line 51: Instead of saying that the severity of LUTS was "related to" hyperglycemia, I would say that the two are "correlated" since correlation is what you actually tested.

Answer: Changes have been made as suggested.

(Abstract section, line 54-55, page 3): “The severity of cn-LUTS was correlated to hyperglycemia”

6. Page 3, line 50: I'm not following your conclusion that this information can be used to "provide novel biomarkers and treatment options." I would say you have overstated your conclusions here.

Answer: Changes have been made as suggested.

The statement (provide novel biomarkers and treatment options.) has been removed from the conclusion subsection of the abstract.

(Abstract section, line 52-54, page 3): A better understanding of urinary microbiota in the development and progression of cn-LUTS in female T2D patients was necessary.

Introduction:

7. Page 4, lines 68-69: Please provide a reference for this statement.

Answer: Changes have been made as suggested.

(Introduction section, line 68-69, page 4): Diabetes is burgeoning worldwide due to an increase in aging population and changes in eating habits [1].


8. Page 4, lines 70-73: I am concerned about the accuracy of your statement that LUTS is a higher risk complication that any other diabetic complication. The reference you provided is specifically focused on diabetic uropathy, so I would like to see a more broad, higher quality reference to support this statement. What about complications like CV disease, obesity, MI, stroke, diabetic neuropathy, foot ulcers, etc? Many of those issues are quite common as well.

Answer: Thank you for your good advice. However, we cannot find the more broad, higher quality reference to support this statement (higher than that of widely recognized complications such as diabetic nephropathy (50%)). Due to cultured negative LUTS receives far less attention, references that focused on other complications like CV disease, obesity, MI, stroke, diabetic neuropathy, foot ulcers did not mention cultured negative LUTS in their articles. Therefore, we prefer to remove the sentence from the introduction section and supplement another reference to support the statement of “80% of individuals with diabetes are accompanied by lower urinary tract symptoms (LUTS)”.

(Introduction section, line 71-72, page 4): 80% of individuals with diabetes are accompanied by lower urinary tract symptoms (LUTS) [1,2].

9. Page 4, lines 68-88: Your "background" section is too long and cumbersome. It is hard to follow the flow of ideas. I think most of the information on page 4 could be summarized in 1-2 sentences. The more important and relevant part of your background is on page 5, so I would focus on that information.

Answer: Changes have been made as suggested. Some contents have been removed from the introduction.

The paucity of knowledge is an important barrier to develop the best prevention and treatment for it. As the completion of the National Institutes of Health Human Microbiome Project, numerous studies suggested that microbial dysbiosis were responsible for promoting disease progression, such as periodontitis, inflammatory bowel disease, colorectal cancer and breast cancer [4-7]. (Introduction section, line 68-79, page 4)

Diabetes is burgeoning worldwide due to an increase in aging population and changes in eating habits [1]. Chronic hyperglycemia is associated with damage, dysfunction, and failure of multiple organ systems, including the genitourinary system. 80% of individuals with diabetes are accompanied by lower urinary tract symptoms (LUTS) [1,2]. Although diabetes-induced LUTS is not life-threatening, it seriously affects patients' life quality. Most diabetic patients are likely suffering from LUTS without evidence of urinary tract infections (UTI), which are refractory and hardly benefit from the conventional treatment. For these patients, it is supposed that the potential pathogenesis of LUTS may be linked to polyuria, oxidative stress and autonomic neuropathy induced by chronic hyperglycemia [3]. However, the underlying pathogenesis of cultured negative LUTS (cn-LUTS) induced by diabetes has not been fully appreciated.

10. Page 5, line 100: You state that increased levels of glucose in the urine changes the "microenvironment of the urinary tract, which might further change the urinary microbiota." What do you mean by "microenvironment," and how does this contrast with "urinary microbiota."

Answer: Thank you for such a good question. Urinary microenvironment means that the environment urothelium contacted with. Urinary microbiota is a key component of the urinary microenvironment. Urine glucose is another important part of the urinary microenvironment, high levels of urine glucose can favor a proper microenvironment to bacterial overgrowth.


11. Page 5, lines 106-109: Your aim needs to be more specific about how you are comparing the urinary microbiota. Specifically, I think you are comparing alpha and beta diversity between diabetic patients with culture-negative LUTS and healthy controls. If that is true, you should state that as your specific aim.

Answer: Changes have been made as suggested.

(Introduction section, line 100-103, page 5): The goal of our study was to characterize urinary microbiota in Chinese T2D females by comparing the alpha and beta diversity as well as the specific genus and to explore the potential pathogenesis of diabetes-induced cn-LUTS from the characteristics of urinary microbiota.
12. Page 5, line 109: You should also clearly state your hypothesis.

Answer: Changes have been made as suggested.

(Introduction section, line 100-103, page 5): The goal of our study was to characterize urinary microbiota in Chinese T2D females by comparing the alpha and beta diversity as well as the specific genus and to explore the potential pathogenesis of diabetes-induced cn-LUTS from the characteristics of urinary microbiota.

Results:

13. Throughout the results section, you include a LOT of information that should be in the methods section. Also, classically the methods section comes after the introduction and before "results." Per the BMC Urology journal guidelines, the methods section should follow the background, and come just before the Results section. Therefore, your manuscript is not formatted properly for this journal. Is there some reason why you put the methods at the end?

Answer: Changes have been made as suggested. We initially submitted our manuscript to BMC Microbiology. Following the journal style of BMC Microbiology, the methods section should follow the conclusion section. However, editor of BMC microbiology considered our manuscript was not suitable for the journal and invited us to use the transfer service to transfer our manuscript to BMC Urology. This is the reason we put the methods in the end. The methods section has been shifted before the results section in the revised manuscript. The information that should be in the methods section has been moved.

(Results section, line 176-277, page 8-13)

14. Page 6, lines 114: The sentence "too little sequencing reads" needs to be explained. How many sequencing reads is too little for analysis? This is information you should provide to your readers, who will mostly be a clinical audience.

Answer: Changes have been made as suggested. The sentence "too little sequencing reads" means we cannot receive raw reads in some samples when the sequence was performed by Illumina Miseq sequencer (Illumina, Inc., USA).

(Results section, line 180-181, page 9): However, 3 samples from patients and 4 from controls were excluded due to the sequencing reads could not be achieved.

15. Page 6, line 114-118: This is not an adequate power calculation. When you do a power (or sample size) calculation, you MUST provide information about what piece of data you are basing your power calculation on, how you expect that data to differ between groups, as well as an estimate of the effect size. You have provided none of those elements, so it's impossible to tell whether you have done your power calculation correctly.

Answer: Thank you for such a good question. There were no studies of urinary microbiota estimated sample size currently due to the lack of data. Our study is cross-sectional with a small sample size. 35 T2D patients and 30 healthy controls were included from June 2017 to December 2017. Finally, statistical analysis and bioinformatics analysis was conducted on data formed 32 T2D patients and 26 healthy controls due to the sequencing results. Therefore, PASS programme was applied to provide
estimates of the power of test. (Results section, line 181-185, page 9): With 32 patients in diabetes group and 26 subjects in the control group (α = 0.05; β = 0.2), we would have 95% power to detect differences at the 0.05 significance level (alpha) using a two-sided Mann-Whitney U test and have 96% power to detect differences at the 0.05 significance level (alpha) using a two-sided Student's t-test.

16. Page 6, lines 119-120: You should describe the baseline characteristics of your study population; don't just say there is "no difference" between groups. More specifically, you should provide a brief summary of the population as a whole, and then values for each of the 2 groups (for example, mean age, parity, a1c, etc). This comment applies to the entire results section.

Answer: Changes have been made as suggested. (Results section, line 186-193, page 9): Except for fasting blood glucose, no significant differences were observed in the demographic and clinical characteristics between diabetes group and control group (for example, age, 56.969±8.014 vs. 57.615±9.239; body mass index, 23.739±4.379 vs. 24.298±3.120; hypertension rate, 14(43.8%) vs. 8(30.8%); etc.). Besides, higher scores of American Urological Association Symptom Index (AUA-SI) (total score, 9.219±6.904 vs. 2.846±3.319; storage score, 5.375±4.612 vs. 1.846±1.617; emptying score, 3.844±4.573 vs. 1.000±2.482) were found in T2D patients (P < 0.05). (Results section, line 198-202, page 10): There were no significant differences in Observed Species (88(11, 245) vs. 135(15, 258)), Chao 1 index (103.60(14.00, 272.00) vs. 149.50(29.00, 339.00)), ACE index 104.30(16.50, 416.70) vs. 146.10(40.70, 428.60), Shannon index (1.70(0.04, 3.48) vs. 1.78(0.01, 3.54)) and Simpson index (0.31(0.06, 0.99) vs. 0.38(0.07,1.00)) between diabetes group and control group (Table1, Figure 1a -1e) (Results section, line 222-228, page 11): Higher Observed Species (Figure 2a, 76(10, 175) vs. 98(54, 234), P = 0.011), Chao1 index (Figure 2b, 83.86(11.00, 210.77) vs. 125.19(62.00, 251.77), P = 0.005), ACE index (Figure 2c, 85.52(13.75, 204.84) vs. 129.82(63.89, 280.30), P = 0.005) were presented in the LS group, while no significant differences were observed in Shannon index (Figure 2d, 1.40(0.04, 3.48) vs. 2.01(0.98, 3.43), P = 0.132) and Simpson index (Figure 2e, 0.44(0.06, 0.99) vs. 0.28(0.07, 0.64), P = 0.202). (Results section, line 254-259, page 12): Higher Shannon index (Figure 4d, 1.369(0.044, 3.478) vs. 2.089(0.979, 3.429), P = 0.033) and lower Simpson index (Figure 4e, 0.461(0.055, 0.987) vs. 0.229(0.069, 0.640), P = 0.029) were presented in LH group, while no significantly difference in Observed Species, Chao1 index and ACE index (Figure 4a, 77(10,204) vs. 86(54,234), P = 0.157; Figure 4b, 97.00(11.00, 213.55) vs. 115.67(63.89,280.30), P = 0.105; Figure 4c, 94.87(0.26, 3.32) vs. 121.30(63.89, 280.30), P = 0.089).

17. Page 7, line 133: Please describe how the urinary microbiota in the diabetic group "clustered away" from the control group. Be more specific. What exactly do you mean?

Answer: Thank you for such a good question. Differences in the composition of urinary microbiota from distinct samples were measured by calculating the distance and shown in PCoA plots. “Urinary microbiota in the diabetic group clustered away from the control group” means that urinary microbiota in different groups clustered in the different area of the PCoA plots, which indicated that the composition of urinary microbiota in the diabetic group differed from that in control group. Changes have been made as suggested.
As shown in Figure 1f – 1h, we found urinary microbiota in diabetes group and control group mainly clustered together in the different area of the PCoA plots, respectively.

18. Page 7, lines 147-148: It is unclear/confusing to title this section "comparison of bioinformatics between T2D patients among different severity of LUTS." I think ALL of the patients in this analysis were diabetic, correct? If that is true, then it's confusing to say you are comparing microbiota "between" diabetics, because that implies you are comparing diabetics to another group (like healthy controls).

Answer: Changes have been made as suggested.

(Results section, line 217-218, page 10): Comparisons of bioinformatics between T2D patients with no to mild LUTS and those with moderate to severe LUTS

19. Page 7, lines 151-152: Why did you dichotomize AUA score? When you dichotomize a continuous variable, you lose much of the variability in our data points and decrease power. Another option would be to model AUA score as a continuous variable.

Answer: Thank you for your excellent point. We want to divide diabetic patients into two groups. One group contained diabetic patients with no to mild LUTS and the other group contained those with moderate to severe LUTS. We tried to explore factors (alternations of urinary microbiota, demographic and clinical characteristics) associated with severe LUTS between the two groups. According to the 2018 EAU guidelines, AUA-SI (IPSS) = 0 was defined as asymptomatic, 1-7 as mild, 8-19 as moderate, and 20-35 as severe. Although AUA-SI (IPSS) was used to apply in evaluating the male LUTS, some studies have found that IPSS could also be applied to assess female LUTS. Thus, it is reasonable to use the AUA-SI score to divide diabetic patients into different groups.

20. Page 8, lines 167-173: Again, please provide information on your raw numbers/values here. Don't just describe in prose what you found. The results section should include numbers and should be organized more clearly.

Answer: We appreciate the reviewer for this constructive suggestion. To identify significantly different bacteria as biomarkers between groups, taxa summaries were reformatted and inputted into LEfSe through the Huttenhower Lab Galaxy Server. The results (plots) of LEfSe only contained relatively increased (decreased) genus without the original number/value. The differentially abundant genera have been visually represented in bar plots if the P-values were statistically significant (α = 0.05) and the absolute value of logarithm of linear discriminant analysis score greater than 2 (Figure 1i, Figure 2i, Figure 4i).

21. Page 9, line 188: This statement is irrelevant to this section of the paper. In this section you are describing the microbial differences between the low a1c and high a1c groups; therefore, it's not relevant that you did not find this correlation between AUA-SI and alpha diversity (that information was already presented in the previous section on lines 157-158).

Answer: We are sorry about this clear mistake. The sentence has been removed from the results section. (Results section, line 261, page 12)
**Discussion:**

22. Somewhere in the discussion section, I think you should comment on why some alpha diversity parameters were different between groups while others were not (ex: the low a1c group had a higher Shannon index but lower Simpson index).

**Answer:** We appreciate the reviewer for this constructive suggestion. I have commented on the alpha diversity before the limitation. In addition, both increasing Shannon index and decreased Simpson index imply a higher alpha diversity (Results section, line 259-261, page 12): Comparison of alpha diversity indicated that significantly decreasing species diversity was discovered in HH cohort. (Discussion section, line 356-368, page 17): In this study, no significant differences in alpha diversity were found between T2D patients and controls but decreased richness as well as species diversity were shown in HS and HH group, respectively. Although a reduction in alpha diversity has been thought as a feature of gastrointestinal diseases, such as ulcerative colitis, Crohn’s disease and colorectal cancer [31-33] as well as an increased vaginal microbiota diversity was associated with bacterial vaginitis [34]. No consistent changes in alpha diversity were found in urologic disorders. Increased alpha diversity was observed in urgency urinary incontinence [11], decreasing alpha diversity was demonstrated in overactive bladder [10], while no significant difference was noted in prostate cancer [35]. In Liu et al.’s study, they reported that urinary microbiota diversity and richness were lower in female T2D patients [14]. Demographic features, exclusive criteria and less microbial species in the urinary tract may be the reasons leading to this contradiction between Liu’s study and ours [14].

23. Page 11, lines 222-242: You don't need to repeat the details of your results in the discussion section. You should provide a short review of the main points of your results, then move on to more of an interpretation or discussion of your results and what they mean for patients/clinicians.

**Answer:** We appreciate the reviewer for this constructive suggestion. Changes have been made as suggested. (Discussion section, line 302-311, page 14-15): Due to the significant difference of HbA1c was found between the LS group and HS group, urinary microbial characteristics in the HH group and LH group were also analyzed. Statistical differences of urinary microbiota composition between HH and LH cohort were found in our analysis (Figure 4f – 4i). In Liu et al.’s study, they found that increased Actinobacteria phylum as well as decreasing Akkermansia muciniphila were positively correlated with fasting blood glucose [14]. It is therefore suggested that poorly controlled blood glucose might further influence the composition of urinary microbiota. Furthermore, we found that most patients in the HH group suffered from more serious LUTS and HbA1c was positively correlated with AUA-SI. These raise the possibility that poorly controlled blood glucose might exacerbate en-LUTS.

24. The discussion section as a whole is cumbersome, wordy and difficult to follow. It should be reformatted to include these main sections:
   a. Briefly summarize your main results
   b. The bulk of the discussion should be your interpretation of the meaning of the results. Compare and contrast your results with other studies.
   c. Describe study strengths and limitations
   d. Provide ideas for future research or next steps.
Answer: Thank you for your considerable advice. The discussion section has been extensively revised following your advice (Discussion section, line 278-395, page 13-18).

25. Page 12, lines 247-256: I'm not following the sequence of ideas in this paragraph, and I can't follow why you would conclude with these current data how Escherichia-Shigella may cause inflammation in the bladder when you did not assess bladder inflammation at all. This conclusion seems outside the scope of the paper.

Answer: Thank you for your good advice. Changes have been made as suggested. (Discussion section, line 317-330, page 15): In both cohorts, one or two genera dominated most of the sequence profiles. Escherichia-Shigella was relatively increased in diabetes group and HS group. Klebsiella and Enterococcus were relatively increased in the diabetes group. Klebsiella is a common conditional pathogenic bacterium that causing respiratory and urinary tract infections when host immunity is impaired [22]. Enterococcus is a nosocomial pathogen which can cause UTI and endocarditis [23]. Increased TNF-α, a proinflammatory cytokine, was discovered in the bladder tissue of T2D mice [24]. Toll-like receptors 4 pathway was also activated in type 1 diabetes rats induced by Streptozotocin [25]. It was a strong indication that inflammation existed in the bladder tissue of diabetic animals. In the present study, diabetic patients that included were suffered from cn-LUTS. As for these patients, standard clinical cultivation procedures have a too low sensitivity to detect the uropathogens. 16S rDNA sequencing of mid-stream urine samples might represent a novel method for diagnosing “UTI” (urinary dysbiosis).

26. Page 12, lines 257-264: Again, this description of the theory of how Campylobacter may be leading to LUTS symptoms is completely outside the scope of your paper and should be removed.

Answer: Thank you for your considerable advice. We apologize for the overstatement of how Campylobacter may be leading to LUTS in the discussion section. The part has been removed.

27. Page 15, line 311: Recommend you remove the comment about "potential harm of catheterization." Urethral catheterization is very low risk and done routinely in Urologic/Urogynecologic practice. Moreover, studies from Loyola University group have demonstrated that catheterized urine specimens for urinary microbiome studies are as good as suprapubic samples, and both are better than clean catch (clean catch urine samples are often contaminated by vaginal flora).

Answer: Thank you for your good advice. The discussion of midstream urine samples was presented in the discussion section. (Discussion section, line 369-377, page 17-18): Owing to the urine samples may be contaminated with microbiota surrounding the urethral orifice, mid-stream urine was collected by the clean catch method with labial separation supervised by the author (Weina Huang). Catheter-derived urine samples were suggested as an alternative method, but the method was considered being invasive and was not ethically workable in patients who do not have a clinical indication for it. Comparison of results from Siddiqui et al.’s urinary microbiota study on female mid-stream urine with the consequences of suprapubic aspirate by Wolfe et al. shown that the major findings were the same [12]. It was a strong indication that results from mid-stream urinary microbiota were reliable.
28. Page 15, lines 314-315: I disagree that future studies should include urodynamic parameters to classify patients into subgroups. Urodynamics are often not accurate; for example, up to 30% of patients with clinical overactive bladder will not have detrusor overactivity on urodynamic testing. Also, which subgroups are you referring to?

Answer: Thank you for such a good question. Changes have been made as suggested. The statement of urodynamic has been removed from the discussion section.

28. Page 15, lines 324-325: I'm not following your conclusion that this information can be used to "provide novel biomarkers and treatment options." You didn't look at biomarkers at all with this study, so you should either remove this conclusion, or provide a justification of how you reached it. In general, I would say you have overstated your conclusions.

Answer: Thank you for such a good advice. LEfSe analysis provided relatively increasing (decreased) genus which could be regarded as biomarkers.

(Methods section, line 167-170, page 8): To identify significantly different bacteria as biomarkers between groups, taxa summaries were reformatted and inputted into Linear discriminant analysis effect size (LEfSe) through the Huttenhower Lab Galaxy Server [20].

(Conclusion section, line 403-406, page 19): A better understanding of urinary microbiota in the development and progression of cn-LUTS in female T2D patients was necessary, which might provide novel diagnostic biomarkers as well as microbiota-targeted therapeutic options.

Methods:

29. Why did you not use expanded urine culture techniques in addition to 16s sequencing? With expanded urine culture, you can get species level information, as well as information about whether the microbes you detect are actually viable.

Answer: Thank you for such a good suggestion. This is a defect that could not be ignored. The expanded urine culture can provide species level information as well as information whether the microbes are actually viable. However, as we know, expanded urine culture techniques only can be performed by several medical centers in the USA recently. Although expanded urine culture can provide more information about urinary microbiota than standard clinical cultivation procedures, the information that expanded urine culture provided is still less 16s sequencing. We need to relatively comprehensively describe the characteristics of the urinary microbiota of T2D patients with cultured negative LUTS in the initial study. Thus, we performed 16s sequencing instead of the expanded urine culture. In the future study, we should add expanded urine culture as a complementary method and make a comprehensive analysis of information that provided by expanded urine culture and 16s sequencing. The statement has been added to the discussion section.

(Discussion section, line 389-393, page 18): Secondly, we cannot confirm that the urinary microbes characterized by 16S rDNA sequencing were actually viable. We should add expanded urine culture as a complementary method and make a comprehensive analysis of information that provided by expanded urine culture and 16S sequencing in the future.

30. Why did you use the AUA symptom index as your questionnaire? Why not use a questionnaire like UDI-6 or OAB-q?
Thank you for such a good question. Diabetes-induced LUTS included early (compensated) stage and advanced (dis compensated) stage. During the early stage, patient present storage symptoms (frequency, urgency, nocturia), however, during the advanced stage, patient present voiding symptoms (urinary retention and urinary incontinence). AUA-SI (IPSS) consists of a total of seven questions that deal with voiding symptoms (incomplete empty, intermittency, weak stream and straining to void) and storage symptoms (frequency, urgency, and nocturia). There have been already some studies found that IPSS can also be applied to assess LUTS in women. Furthermore, UDI-6 and OAB-q have not been widely used in China.

31. Please explain your rational for using parametric testing for comparison of baseline characteristics, but non-parametric testing (Mann Whitney U) for the bioinformatics measures. Were your baseline characteristic data normally distributed? Did you test for normality? I would suspect that your data (baseline or demographic data) are normally distributed given the small sample sizes. If they are not, then you should be using non-parametric tests for those comparisons as well.

Answer: We are sorry about this clear mistake. Changes have been made as suggested. Bioinformatics data were not normally distributed, thus, non-parametric testing (Mann-Whitney U) was applied for bioinformatics measures. Results of tests of normality for characteristic data were shown below. We found that some characteristic data were not normally distributed, so corrective non-parametric tests were applied for these data. Table 1, Table 2 and Table S2 have been corrected (page 26-27).

Tests of Normality for diabetes group

<table>
<thead>
<tr>
<th>Tests of Normality</th>
<th>Kolmogorov-Smirnova</th>
<th>Shapiro-Wilk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Statistic</td>
<td>df</td>
</tr>
<tr>
<td>Age</td>
<td>.114</td>
<td>32</td>
</tr>
<tr>
<td>BMI</td>
<td>.152</td>
<td>32</td>
</tr>
<tr>
<td>FBG</td>
<td>.124</td>
<td>32</td>
</tr>
<tr>
<td>Creatinine</td>
<td>.346</td>
<td>32</td>
</tr>
<tr>
<td>eGFR</td>
<td>.100</td>
<td>32</td>
</tr>
<tr>
<td>AUAT</td>
<td>.136</td>
<td>32</td>
</tr>
<tr>
<td>AUAS</td>
<td>.197</td>
<td>32</td>
</tr>
<tr>
<td>AUSV</td>
<td>.200</td>
<td>32</td>
</tr>
</tbody>
</table>

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Tests of Normality for control group

<table>
<thead>
<tr>
<th>Tests of Normality</th>
<th>Kolmogorov-Smirnova</th>
<th>Shapiro-Wilk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Statistic</td>
<td>df</td>
</tr>
<tr>
<td>Age</td>
<td>.178</td>
<td>26</td>
</tr>
<tr>
<td>BMI</td>
<td>.172</td>
<td>26</td>
</tr>
<tr>
<td>FBG</td>
<td>.160</td>
<td>26</td>
</tr>
<tr>
<td>Creatinine</td>
<td>.123</td>
<td>26</td>
</tr>
<tr>
<td>eGFR</td>
<td>.076</td>
<td>26</td>
</tr>
<tr>
<td>AUAT</td>
<td>.254</td>
<td>26</td>
</tr>
<tr>
<td>AUAS</td>
<td>.200</td>
<td>26</td>
</tr>
<tr>
<td>AUSV</td>
<td>.426</td>
<td>26</td>
</tr>
</tbody>
</table>
Tests of Normality for HS group

<table>
<thead>
<tr>
<th>Tests of Normality</th>
<th>Kolmogorov-Smirnov</th>
<th>Shapiro-Wilk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistic</td>
<td>df</td>
<td>Sig.</td>
</tr>
<tr>
<td>Age</td>
<td>.174</td>
<td>17</td>
</tr>
<tr>
<td>BMI</td>
<td>.172</td>
<td>17</td>
</tr>
<tr>
<td>FBG</td>
<td>.152</td>
<td>17</td>
</tr>
<tr>
<td>Creatinine</td>
<td>.407</td>
<td>17</td>
</tr>
<tr>
<td>eGFR</td>
<td>.138</td>
<td>17</td>
</tr>
<tr>
<td>AUAT</td>
<td>.153</td>
<td>17</td>
</tr>
<tr>
<td>AUAS</td>
<td>.130</td>
<td>17</td>
</tr>
<tr>
<td>AUAV</td>
<td>.191</td>
<td>17</td>
</tr>
</tbody>
</table>

* This is a lower bound of the true significance.
a. Lilliefors Significance Correction

Tests of Normality for LS group

<table>
<thead>
<tr>
<th>Tests of Normality</th>
<th>Kolmogorov-Smirnov</th>
<th>Shapiro-Wilk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistic</td>
<td>df</td>
<td>Sig.</td>
</tr>
<tr>
<td>Age</td>
<td>.155</td>
<td>15</td>
</tr>
<tr>
<td>BMI</td>
<td>.151</td>
<td>15</td>
</tr>
<tr>
<td>FBG</td>
<td>.157</td>
<td>15</td>
</tr>
<tr>
<td>Creatinine</td>
<td>.226</td>
<td>15</td>
</tr>
<tr>
<td>eGFR</td>
<td>.103</td>
<td>15</td>
</tr>
<tr>
<td>AUAT</td>
<td>.221</td>
<td>15</td>
</tr>
<tr>
<td>AUAS</td>
<td>.215</td>
<td>15</td>
</tr>
<tr>
<td>AUAV</td>
<td>.328</td>
<td>15</td>
</tr>
</tbody>
</table>

* This is a lower bound of the true significance.
a. Lilliefors Significance Correction

Tests of Normality for HH group

<table>
<thead>
<tr>
<th>Tests of Normality</th>
<th>Kolmogorov-Smirnov</th>
<th>Shapiro-Wilk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistic</td>
<td>df</td>
<td>Sig.</td>
</tr>
<tr>
<td>Age</td>
<td>.160</td>
<td>17</td>
</tr>
<tr>
<td>BMI</td>
<td>.150</td>
<td>17</td>
</tr>
<tr>
<td>FBG</td>
<td>.190</td>
<td>17</td>
</tr>
<tr>
<td>Creatinine</td>
<td>.462</td>
<td>17</td>
</tr>
<tr>
<td>eGFR</td>
<td>.169</td>
<td>17</td>
</tr>
<tr>
<td>AUAT</td>
<td>.116</td>
<td>17</td>
</tr>
<tr>
<td>AUAS</td>
<td>.122</td>
<td>17</td>
</tr>
<tr>
<td>AUAV</td>
<td>.224</td>
<td>17</td>
</tr>
</tbody>
</table>

* This is a lower bound of the true significance.
a. Lilliefors Significance Correction

Tests of Normality for LH group

<table>
<thead>
<tr>
<th>Tests of Normality</th>
<th>Kolmogorov-Smirnov</th>
<th>Shapiro-Wilk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistic</td>
<td>df</td>
<td>Sig.</td>
</tr>
<tr>
<td>Age</td>
<td>.160</td>
<td>17</td>
</tr>
<tr>
<td>BMI</td>
<td>.150</td>
<td>17</td>
</tr>
<tr>
<td>FBG</td>
<td>.190</td>
<td>17</td>
</tr>
<tr>
<td>Creatinine</td>
<td>.462</td>
<td>17</td>
</tr>
<tr>
<td>eGFR</td>
<td>.169</td>
<td>17</td>
</tr>
<tr>
<td>AUAT</td>
<td>.116</td>
<td>17</td>
</tr>
<tr>
<td>AUAS</td>
<td>.122</td>
<td>17</td>
</tr>
<tr>
<td>AUAV</td>
<td>.224</td>
<td>17</td>
</tr>
</tbody>
</table>

* This is a lower bound of the true significance.
a. Lilliefors Significance Correction
<table>
<thead>
<tr>
<th>Statistic</th>
<th>df</th>
<th>Sig.</th>
<th>Statistic</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>.157</td>
<td>15</td>
<td>.200*</td>
<td>.914</td>
<td>15</td>
</tr>
<tr>
<td>BMI</td>
<td>.166</td>
<td>15</td>
<td>.200*</td>
<td>.898</td>
<td>15</td>
</tr>
<tr>
<td>FBG</td>
<td>.207</td>
<td>15</td>
<td>.082</td>
<td>.849</td>
<td>15</td>
</tr>
<tr>
<td>Creatinine</td>
<td>.197</td>
<td>15</td>
<td>.122</td>
<td>.826</td>
<td>15</td>
</tr>
<tr>
<td>eGFR</td>
<td>.107</td>
<td>15</td>
<td>.200*</td>
<td>.965</td>
<td>15</td>
</tr>
<tr>
<td>AUAT</td>
<td>.222</td>
<td>15</td>
<td>.045</td>
<td>.835</td>
<td>15</td>
</tr>
<tr>
<td>AUAS</td>
<td>.229</td>
<td>15</td>
<td>.034</td>
<td>.838</td>
<td>15</td>
</tr>
<tr>
<td>AUAV</td>
<td>.283</td>
<td>15</td>
<td>.002</td>
<td>.727</td>
<td>15</td>
</tr>
</tbody>
</table>

* This is a lower bound of the true significance.

a. Lilliefors Significance Correction

32. Page 18: Why do you refer to the Mann Whitney U test in lines 380-381, but to the "Wilcoxon rank test" on line 391. These are the same test, so you should use a consistent term to describe them.

Answer: Changes have been made as suggested.

(Methods section, line 173, page 8): Mann-Whitney U test

33. Page 18, line 382: You should spell out "MRPP" the first time you use it, and also explain to your readers (many of whom will be primarily a clinical audience) what MRPP is and how it is used.

Answer: Changes have been made as suggested.

(Methods section, line 160-163, page 8): Multiple Response Permutation Procedure (MRPP, a nonparametric test which applied to test beta diversity values between groups) was performed to test the differences of Bray Curtis, weighted UniFrac, unweighted UniFrac distance metrics between groups in QIIME.

34. In general, I think the methods section is far too vague - you should clearly describe each step of the methods so that it can be clearly reproduced by other investigators. I don't think your study is easily reproducible due to the lack of detail in the "methods" section. You could add supplemental material to provide more details of your methods.

Answer: Changes have been made as suggested. The bioinformatics analysis of methods section has been rewritten and more detailed information has been supplemented in it.

(Methods section, line 144-175, page 7-8):

Bioinformatics analysis

16S rDNA sequence data were analyzed using the Quantitative Insights Into Microbial Ecology (QIIME 1.80) [15]. Sequences were clustered into individual operational taxonomic units (OTUs) at a default similarity level of 97% using USEARCH 7.0.1090 software [16], and subsequently, chimera detection was performed using the UCHIME method [17]. A single representative sequence from each clustered OTU was used to align to the Silva and Greengenes database using Ribosomal Database Project Classifier [18].

Alpha diversity indexes including the Observed Species, Chao 1 index, ACE index, Shannon index, and Simpson index were calculated by mothur 1.31.2 software, with the Mann-Whitney U test used to analyze differences by R 3.0.3 software. Observed Species, Chao1 index, and ACE index represented bacterial richness, while Shannon index and Simpson index were quantitative measures of bacterial
diversity that reflecting both species richness and evenness.

To compare microbial composition between groups, Principal coordinate analysis (PCoA) was applied on Bray Curtis, weighted UniFrac and unweighted UniFrac distance metrics to generate three-dimensional plots in QIIME. Multiple Response Permutation Procedure (MRPP, a nonparametric test which applied to test beta diversity values between groups) was performed to test the differences of Bray Curtis, weighted UniFrac, unweighted UniFrac distance metrics between groups in QIIME. Significant bacteria on the relative abundance was tested by Metastats Test. FDR (false discovery rate) was the corrected p-value, and *FDR < 0.05 was considered significant for tests. The p-value was corrected by p.adjust in the R with the method of Benjamini-Chochberg [19]. To identify significantly different bacteria as biomarkers between groups, taxa summaries were reformatted and inputted into Linear discriminant analysis effect size (LEfSe) through the Huttenhower Lab Galaxy Server [20]. In the settings of LEfSe, the Kruskal-Wallis sum rank test (ɑ = 0.05) was used to detect taxa with significant differential abundance at first; as a second step, biological consistency was then investigated with a set of pairwise tests among subclasses using the Mann-Whitney U test; finally, linear discriminant analysis (LDA) was used to estimate the effect size of differentially abundant genera. The threshold on the logarithmic LDA score for discriminative features was 2.

35. Page 18, lines 395-396: Why did you use the threshold of 2 for the linear discriminant analysis? You should provide a rationale for this threshold and what it means for the reader.

Answer: Thank you for such a good question. LEfSe was performed via the Huttenhower Lab Galaxy Server (http://huttenhower.sph.harvard.edu/galaxy/). During the analysis procedure, the logarithmic linear discriminant analysis (LDA) score threshold of 2 was a default value. In addition, a threshold of 2 was widely used in the LEfSe analysis.

Tables/Figures:
* Table 1 and Table 2: please specify in the footnote whether any data were missing.

Answer: Changes have been made as suggested.
(Table section, line 544, page 26): ※No data were missing.
(Table section, line 549, page 27): ※No data were missing.

* Figures 1, 2, and 4: The sub-figures (f-h) are way too small and unclear to see. Therefore, I'm unable to assess the quality of these figures, or interpret what they are showing.

Answer: Thank you for such a good advice. The figures in the pdf document were compressed, which may cause the quality of the figure decreasing. We will re-upload all the figures presented in the manuscript in tiff format as attachments.

* Table 1 and 2 are duplicated in the manuscript. You can remove the duplicates.

Answer: Thank you for proposing such a good question. Table 1 and Table 2 are not duplicated. Comparisons of demographic and clinical characteristics as well as alpha diversity between diabetes group and control group were shown in Table 1. In addition, Comparisons of demographic and clinical characteristics as well as alpha diversity between the HS group and LS group were shown in Table 2.
* Figures 1-4: I appreciate the separate page of figure legends, but I think each figure should still include the title on the same page as the figure. It was difficult to tell which figure was #1, which was figure #2, etc.

Answer: Thank you for such a good suggestion. We will add the title on the same page as the figure.

Noor Peter N. Buchholz (Reviewer 2):
The epidemic boom of diabetes becomes a significant challenge in everyday urology work. Over 80 % of diabetic patients suffer from LUTS with poorly understood multifactorial etiology, which are notoriously hard to treat with the standard therapeutic regiments and approaches.

The authors present a very interesting study, aiming to present and assess a novel point of view in the pathoetiology of diabetic-induced LUTS - the urinary tract dysbiosis. The authors aim their research on the concept of urinary microbiota changes, associated with cultured negative LUTS in female T2D patients. They use a sophisticated genetic and statistical method to assess the changes in urinary microbiota in T2D and their correlation with diabetic LUTS, which may give new pathways for diagnosis and treatment in these patents.

The methodology of the authors is according to standards regarding data extraction and analysis, quality assessment of the data and statistical analysis. The design of the study is very sophisticated and extensive, including differences in urinary microbiota between controls and T2D patients, between T2D patients with mild and severe LUTS, and between T2D patients with good or poor diabetic control. In all these cases authors find significant differences which represent the basis for a future research in this entirely new field.

The limitations of this study are acknowledged and discussed by authors - most importantly the very small number of patients for a problem of such complexity, but quite enough for a pilot study which could initiate serious further development of entire new concept and treatment of LUTS. The other limitations which are discussed (patient heterogeneity, urine sampling, AUA-SI usage instead of urodynamics) in the reviewer opinion had no negative impact on the concept of the study. Especially the implementation of urodynamic study and at a lesser degree urine sampling through catheterization, will make larger scale study close to impossible, and the result not applicable to everyday practice. All of these problems should be addressed in future multi-centered prospective studies which includes them in their protocols.

The conclusions are adequately substantiated and represent a comprehensive basic science study and a solid base to further investigation on the subject.

The references are appropriate and relevant to the subject and my recommendation is to accept this manuscript for publication.

Thank you very much for your comments concerning our manuscript.

Andrea Ticinesi (Reviewer 3):
This is an interesting study exploring a relevant research question (whether LUTS with negative cultures in diabetic patients may be associated with a distinct urinary microbiota composition).
The investigation of urinary microbiota is intriguing and brings some important innovations in the field of functional urogynecologic disorders. The paper is overall well written and clear, although some specific elements of writing style could be substantially improved. The findings are not conclusive, but suggest a link between urinary microbiota dysbiosis and LUTS severity, that should be investigated in future, larger studies with a prospective design.

I have the following comments:

1. The authors declare that they adopted "16S rDNA" analysis. I guess they intended 16S rRNA microbial profiling (i.e. metagenomics). More in general, the microbiological analysis procedures should be better explained, especially for an audience who is not friendly with that complex matter.

   Answer: Thank you for such a good suggestion. In our study, we performed 16S rDNA sequencing (16S rRNA gene sequencing) to identify the urinary microbiota. 16S rDNA sequencing differs from Shotgun sequencing (i.e. metagenomics). During sequencing procedures, firstly, DNA extraction was performed using the cultured cells protocol supplied with the DNeasy Blood and Tissue Kit (Qiagen, Germany). Then, specific primer sets for V3-V4 regions were selected to perform PCR amplification of 16S rDNA. Hence, I think this sequencing method should be called 16S rDNA sequencing.

   The bioinformatics analysis of methods section has been rewritten and more detailed information has been supplemented in it.

   (Methods section, line 144-175, page 7-8):
   Bioinformatics analysis
   16S rDNA sequence data were analyzed using the Quantitative Insights Into Microbial Ecology (QIIME 1.80) [15]. Sequences were clustered into individual operational taxonomic units (OTUs) at a default similarity level of 97% using USEARCH 7.0.1090 software [16], and subsequently, chimera detection was performed using the UCHIME method [17]. A single representative sequence from each clustered OTU was used to align to the Silva and Greengenes database using Ribosomal Database Project Classifier [18].

   Alpha diversity indexes including the Observed Species, Chao 1 index, ACE index, Shannon index, and Simpson index were calculated by mothur 1.31.2 software, with the Mann-Whitney U test used to analyze differences by R 3.0.3 software. Observed Species, Chao1 index, and ACE index represented bacterial richness, while Shannon index and Simpson index were quantitative measures of bacterial diversity that reflecting both species richness and evenness.

   To compare microbial composition between groups, Principal coordinate analysis (PCoA) was applied on Bray Curtis, weighted UniFrac and unweighted UniFrac distance metrics to generate three-dimensional plots in QIIME. Multiple Response Permutation Procedure (MRPP, a nonparametric test which applied to test beta diversity values between groups) was performed to test the differences of Bray Curtis, weighted UniFrac, unweighted UniFrac distance metrics between groups in QIIME.

   Significant bacteria on the relative abundance was tested by Metastats Test. FDR (false discovery rate) was the corrected p-value, and *FDR < 0.05 was considered significant for tests. The p-value was corrected by p.adjust in the R with the method of Benjamini-Chochberg [19]. To identify significantly different bacteria as biomarkers between groups, taxa summaries were reformatted and inputted into Linear discriminant analysis effect size (LEfSe) through the Huttenhower Lab Galaxy Server [20]. In the settings of LEfSe, the Kruskal-Wallis sum rank test (α = 0.05) was used to detect taxa with significant differential abundance at first; as a second step, biological consistency was then investigated with a set of pairwise tests among subclasses using the Mann-Whitney U test; finally, linear discriminant analysis (LDA) was used to estimate the effect size of differentially abundant genera. The threshold on the logarithmic LDA score for discriminative features was 2.
2. Metagenomics is a well-established technique for the analysis of fecal microbiota. However, only a few studies have been performed on urinary samples so far. Urine samples may be technically difficult to analyze with metagenomics, due to low concentration of bacteria and high risk of contamination of genital skin microbiota. Of course, catheterization would have reduced the risk of contamination, but this is not ethically feasible in patients who do not have a clinical indication for it. Are there some data in the literature supporting the reliability of assessing midstream urinary microbiome instead of catheter-derived urinary microbiome? I encourage the authors to better explain how they dealt with these issues in their experience in both the methods and discussion section. This is fundamental for a correct interpretation of results.

Answer: thank you for your suggestion. These issues have been supplemented in the discussion section and method section.

(Methods section, line 119-121, page 6): Mid-stream urine (50ml) was collected with the labial separation which was supervised by the author (Weina Huang) and 20ml was left for standard cultivation to exclude UTI.

(Discussion section, line 369-377, page 17-18): Owing to the urine samples may be contaminated with microbiota surrounding the urethral orifice, mid-stream urine was collected by the clean catch method with labial separation supervised by the author (Weina Huang). Catheter-derived urine samples were suggested as an alternative method, but the method was considered being invasive and was not ethically workable in patients who do not have a clinical indication for it. Comparison of results from Siddiqui et al.’s urinary microbiota study on female mid-stream urine with the consequences of suprapubic aspirate by Wolfe et al. shown that the major findings were the same [12]. It was a strong indication that results from mid-stream urinary microbiota were reliable.

3. I have some concerns about the methodology of sample size calculation, and invite the authors to provide better justification for it. Which was the primary endpoint of the study? Was the calculation based on it? Maybe the work by Kelly BJ et al. (Bioinformatics 2015) can be useful to authors to better define the sample size.

Answer: Thank you for such a good question. There were no studies of urinary microbiota estimated sample size currently due to the lack of data. Our study is cross-sectional with a small sample size. 35 T2D patients and 30 healthy controls were included from June 2017 to December 2017. Finally, statistical analysis and bioinformatics analysis was conducted on data formed 32 T2D patients and 26 healthy controls due to the sequencing results. Therefore, PASS programme (only need to input the sample size) were applied to provide estimates of the power of test.

(Results section, line 181-185, page 9): With 32 patients in diabetes group and 26 subjects in the control group (α = 0.05; β = 0.2), we would have 95% power to detect differences at the 0.05 significance level (alpha) using a two-sided Mann-Whitney U test and have 96% power to detect differences at the 0.05 significance level (alpha) using a two-sided Student's t-test.

4. How were the diabetic patients treated? Anti-diabetic drugs - and particularly metformin - have a well-known effect on the gut microbiota composition. Moreover, a large number of drugs may influence gut microbiota composition, and their effect could be present even in the urinary microbiota.

Answer: Thank you for your question. The content has been added to the discussion section.  

(Discussion section, line 378-384, page 18): Participants who used antibiotics recently have been excluded, however, numerous drugs may influence gut microbiota composition, and their effect may be
present even in the urinary microbiota. Anti-diabetic drugs, particularly metformin, also have a well-known effect on the gut microbiota composition [36]. Although all the T2D patients included in our study were hospitalized and most of them were treated with insulin (data not shown), we could not fully avoid the interference of anti-diabetic drugs. The correlation between urinary microbiota and anti-diabetic drugs should be studied.

5. The authors should make an hypothesis on the origin of the microbial populations they detected in urine samples: were they an autoctonous population or were they derived from the intestinal microbiome?

Answer: Thank you for your question. Although the gut microenvironment differs from the urinary tract, some microbial populations that derived from gut may adapt urinary tract microenvironment. We can’t fully exclude the possibility that urinary microbiota deriving from the intestinal microbiota. However, Thomas-White K et al. found that there were significant differences between gut microbiota and urinary microbiota. It was a strong indication that urinary microbiota should be an autochthonous population.


6. The discussion section is too long and should be extensively revised. I suggest to include only a brief summary of results (much information is duplicated from the results section), and to focus on interpretation.

Answer: Thank you for such a good suggestion. The discussion section has been extensively revised.

(Discussion section, line 278-395, page 13-18)

7. Some data contained in the abstract are confusing. I suggest to rephrase the sentence on alpha-diversity (lines 32-36), because it is not clear in which comparison the significant differences were found.

Answer: Thank you for your suggestion. Changes have been made as suggested.

(Abstract section, line 35-41, page 2): However, statistically decreased richness (ACE index and Chao 1 index, 85.52(13.75, 204.84) vs. 129.82(63.89, 280.30) and 83.86(11.00, 210.77) vs. 125.19(62.00, 251.77), P=0.005; Observed Species, 76(10, 175) vs. 98(54, 234), P=0.011) and decreased species diversity (Shannon index, 1.37(0.04, 3.48) vs. 2.09(0.98, 3.43), P = 0.033; Simpson index, 0.46 (0.06, 0.99) vs. 0.23(0.07, 0.64), P = 0.029) were shown in moderate-to-severe LUTS group and high Hemoglobin A1c group, respectively.

8. In line 345 the authors state that every participant to the study provided a urine sample for cultures. The negativity of cultures should be emphasized also in the results. The content will be presented in the results section.

Answer: Changes have been made as suggested.

(Results section, line 179-180, page 9): All the urine samples from these participants were verified cultured negative during the standard culture procedure.
9. Could it be possible that patients with severe LUTS have a Escherichia coli infection that cannot be detected with traditional urine cultures (that is, urine cultures have too low sensitivity to detect these infections)? If it was the case, metagenomics of midstream urine samples could represent a novel method for diagnosing UTIs in this category of patients. The authors may include some perspectives on that in the discussion section.

Answer: Changes have been made as suggested.

(Discussion section, line 317-330, page 15): In both cohorts, one or two genera dominated most of the sequence profiles. Escherichia-Shigella was relatively increased in diabetes group and HS group. Klebsiella and Enterococcus were relatively increased in the diabetes group. Klebsiella is a common conditional pathogenic bacterium that causing respiratory and urinary tract infections when host immunity is impaired [22]. Enterococcus is a nosocomial pathogen which can cause UTI and endocarditis [23]. Increased TNF-α, a proinflammatory cytokine, was discovered in the bladder tissue of T2D mice [24]. Toll-like receptors 4 pathway was also activated in type 1 diabetes rats induced by Streptozotocin [25]. It was a strong indication that inflammation existed in the bladder tissue of diabetic animals. In the present study, diabetic patients that included were suffered from cn-LUTS. As for these patients, standard clinical cultivation procedures have a too low sensitivity to detect the uropathogens. 16S rDNA sequencing of mid-stream urine samples might represent a novel method for diagnosing “UTI” (urinary dysbiosis).

10. Careful English language revision is needed throughout the paper.

Answer: Thank you for your suggestion. Careful English language revision was performed by our Colleague who was skillful with English.

RM Sathish Kumar, MCh (Reviewer 4):
It is with interest I read the paper on dysbiosis. Seems to have initiated new thought process on LUTS pathogenesis. Asymptomatic bacteriruria causing LUTS is explained in the paper.

1. The statistical process may be explained in an easy method.

Answer: Thank you for such a good suggestion. The bioinformatics analysis of methods section has been rewritten and more detailed information has been supplemented in it.

(Methods section, line 144-175, page 7-8):
Bioinformatics analysis
16S rDNA sequence data were analyzed using the Quantitative Insights Into Microbial Ecology (QIIME 1.80) [15]. Sequences were clustered into individual operational taxonomic units (OTUs) at a default similarity level of 97% using USEARCH 7.0.1090 software [16], and subsequently, chimera detection was performed using the UCHIME method [17]. A single representative sequence from each clustered OTU was used to align to the Silva and Greengenes database using Ribosomal Database Project Classifier [18].

Alpha diversity indexes including the Observed Species, Chao 1 index, ACE index, Shannon index, and Simpson index were calculated by mothur 1.31.2 software, with the Mann-Whitney U test used to analyze differences by R 3.0.3 software. Observed Species, Chao1 index, and ACE index represented
bacterial richness, while Shannon index and Simpson index were quantitative measures of bacterial diversity that reflecting both species richness and evenness.

To compare microbial composition between groups, Principal coordinate analysis (PCoA) was applied on Bray Curtis, weighted UniFrac and unweighted UniFrac distance metrics to generate three-dimensional plots in QIIME. Multiple Response Permutation Procedure (MRPP, a nonparametric test which applied to test beta diversity values between groups) was performed to test the differences of Bray Curtis, weighted UniFrac, unweighted UniFrac distance metrics between groups in QIIME.

Significant bacteria on the relative abundance was tested by Metastats Test. FDR (false discovery rate) was the corrected p-value, and *FDR < 0.05 was considered significant for tests. The p-value was corrected by p.adjust in the R with the method of Benjamini-Chochberg [19]. To identify significantly different bacteria as biomarkers between groups, taxa summaries were reformatted and inputted into Linear discriminant analysis effect size (LEfSe) through the Huttenhower Lab Galaxy Server [20]. In the settings of LEfSe, the Kruskal-Wallis sum rank test (α = 0.05) was used to detect taxa with significant differential abundance at first; as a second step, biological consistency was then investigated with a set of pairwise tests among subclasses using the Mann-Whitney U test; finally, linear discriminant analysis (LDA) was used to estimate the effect size of differentially abundant genera. The threshold on the logarithmic LDA score for discriminative features was 2.