Author’s response to reviews

Title: Diversity of HIV-1 genotypes and high prevalence of pretreatment drug resistance in newly diagnosed HIV-infected patients in Shanghai, China

Authors:

Zhen-yan Wang (a1s2r@163.com)
Min Zhang (zhangminlisa@126.com)
Ren-fang Zhang (zhangrenfang@shphc.org.cn)
Li Liu (liulishaphc@163.com)
Yin-zhong Shen (shenyinzhong@shphc.org.cn)
Jiang-rong Wang (wangjiangrong@shphc.org.cn)
Hong-zhou Lu (luhongzhou@fudan.edu.cn)

Version: 1 Date: 31 Jan 2019

Author’s response to reviews:

Dear Editor,

We are grateful for the opportunity to revise our manuscript for publication in the BMC Infectious Diseases. Please see a detailed response to the comments from the reviewers below. We hope that you find the manuscript improved and acceptable for publication.

With regards,

Hongzhou Lu (on behalf of the authors)

Fudan University, Shanghai Public Health Clinical Center
2901 Caolang Road, Jinshan district, Shanghai China

Response to comments from Associate Editor:

Thank you for the detailed comments on the manuscript. In our study, we did find that the prevalence of PDR was very high among the newly diagnosed AIDS patients in Shanghai. However, for nearly half of the HIV-1 strains with mutation, the degree of resistance was
classified as potential low level. This finding was highlighted both in the title and in the abstract, and also be discussed.

Major points:


Response: According to the new guideline, we use pretreatment drug resistance (PDR) instead of transmitted drug resistance (TDR) to describe HIV drug resistance in our study.

2. The conclusion that the HIV-1 distribution in Shanghai is “extremely” diverse (consisting of five main subtypes/CRFs and seven minor variants) may be misleading. It is diverse, for sure, but this level of diversity is very common. In fact, many countries and regions around the world display a much more diverse HIV-1 distribution than the one reported here. Consider to remove the use of “extremely”.

In this context, it would also be very interesting for the reader to know more about the HIV-epidemic in Shanghai and the region, if such data is available? Prevalence/incidence etc.

Response: We removed the word “extremely” for description of HIV-1 distribution in Shanghai in the conclusion part.

3. Both the introduction and discussion sections are missing references in several places to support the written statements. For example, the statement that HIV-1 genotypes vary remarkable in transmission route is questionable and need to be supported by proper referencing.

Response: We have amended the wording and indicated the reference in the text.
“…HIV-1 genotypes are associated with transmission routes, and vary in epidemic size and distribution features [6]…”

4. It is stated that "with the development of social economy and the rapid growth of migrants, the epidemiological pattern of HIV/AIDS may have changed". This is a very vague statement. What did it change from? Also, it would be helpful to exemplify what the authors mean by "epidemiological pattern".

Response: we rewrote this paragraph as follows:
“…Moreover, with the development of social economy and the rapid growth of migrants, the distribution of HIV-1 genotypes may have become more diverse and complex due to HIV-1 population movements and the ongoing recombination between different HIV-1 subtypes…”

5. P. 6, lines 120-121: Inclusion criteria 1 is not clear - what does newly diagnosed within 3 months mean? Please clarify.

Response: We have amended the inclusion criteria 1 as follows:

1) visited SPHCC during the period from Jan 1st, 2017 to Nov 30th, 2017, and had been diagnosed with HIV/AIDS within three months before the visit.

6. P.7. The PCR protocol and primers used. Is this an established (previously published) protocol, or was this optimized for this study? Please clarify and if previously published/used, indicate a reference.

Response: The PCR protocol and primers used in our study were as described previously, and we have indicated the reference in the revised manuscript (Song YX, Xin RL, Li ZC, Yu HW, Lun WH, Ye J, et al. Prevalence of transmitted drug resistance among HIV-1 treatment-naive patients in Beijing. Epidemiol Infect. 2018; 146:339-44.)

7. Starting on line 138: The state-of-the-art is to verify the HIV subtype/CRF using phylogenetics with reference to the Los Alamos Sequence Database reference dataset (preferably maximum-likelihood or Bayesian with a robust branch support method). A BLAST approach (which is poorly explained in the report) is not sufficient for proper subtype/CRF assignment. A similarity scan, may be used as an explorative analysis, for example COMET. However, these algorithms suffer from low sensitivity of assigning correct CRF classification. This part of the study needs to be entirely redone with proper methodology and reporting.

Response: We have reanalyzed and identified HIV-1 genotypes using the proper methodology - phylogenetic analysis. HIV-1 pol sequences, together with reference sequences of different subtypes and CRFs, were aligned and further edited manually using the BIOEDIT version 7 (www.mbio.ncsu.edu/BioEdit/bioedit.html). All the subtyping reference sequences were downloaded from the Los Alamos HIV database (https://www.hiv.lanl.gov/content/sequence/HIV/mainpage.html). The phylogenetic trees were generated using the Neighbor-joining method. We added the description of methodology and the results in the revised manuscript accordingly.

8. Please add definitions for the degree of resistance on p. 7, lines 140-142 and references (if available). Which algorithm was used? The calibrated population resistance (CPR) tool is recommended for scoring pretreatment drug resistance? Please clarify.
Response: DRMs and resistance levels were determined based on Stanford University HIV Drug Resistance Database (HIVDB): HIVdb Program (https://hivdb.stanford.edu/hivdb/by-sequences/). The degree of drug resistance to each ARV was divided into five levels: susceptible, potential low-level resistance, low-level resistance, intermediate resistance, and high-level resistance, according to the HIVDB Genotypic Resistance Test (GRT) Interpretation System (Updated October 2018, https://hivdb.stanford.edu/assets/media/genotypic-resistance-test-interpretation-system-oct2018.57d61710.pdf).

9. P. 7, lines 145-146: How was normality assessed?

Response: Normality of data was assessed by the Kolmogorov–Smirnov test. We have added the corresponding content in the method section.

10. Another point is the CD4 status of the newly diagnosed participants. A median CD4 count of 275 indicates that the majority of participants were very immunosuppressed and they have presumably been infected for a relatively long time. How does this affect the interpretation of the data and reflection of the current HIV epidemic in Shanghai?

Response: The patients enrolled in our study were newly diagnosed with HIV-1 infection in the year 2017. The “new diagnosis” does not mean “the recent/current infection”. In fact, many HIV-1 infected patients didn’t know when they had got HIV-1 infection. Some participants might have been infected for a relatively long time, and just knew the status of HIV-1 infection recently. We acknowledge that such case may affect the real reflection of current HIV epidemic, and added the relevant discussion about this.

11. P. 8, line176: Check reporting of the numbers and percentages, e.g. 4 of 7 cannot equal 8.5%?

Response: We are sorry to have made a mistake and have corrected the number.

12. Multiple comparisons have been made. Was correction for multiple testing done? If not, motivate why this was not necessary.

Response: We have performed multinomial logistic regression analysis for risk factors of pretreatment drug resistance (Table 3).

13. P. 9, lines 193-196 (also in abstract): This paragraph describes the results from a Chi2 test comparing the distribution between three groups. The author interpret this as subtype B having higher proportions of TDR mutations than CRF01_AE and CRF07_BC. This is not correct! The test only asses the overall difference in distributions between the three groups and does not say anything about which specific groups that differ. This must be tested by pairwise testing (and
correction for multiple testing, e.g. Bonferroni). Inspection of the effect estimates suggest that it is the CRF07_BC that is lower than the other two groups. However, this must be verified by pairwise comparisons.

Response: We had compared the distribution of HIV drug resistance mutations between three groups using an improper statistical method and offered an inappropriate explanation of the result. Therefore, we removed this part in the revised manuscript.

14. P. 10, lines 202-216: This paragraph is very descriptive. It would be informative to see some statistics and comparisons in this section.

Response: We performed statistical analysis in this part; we compared the prevalence of HIV-1 PDR to different ARV drugs using Chi-square test.

15. P. 11, lines 221-222: Mutation rates have not been assessed in this study. Please correct.

Response: The “mutation rate” was corrected as “mutation frequency”.

16. P. 11, lines 222-223: "...half of the mutated virus strains...". Mutated from what? Do you mean “half of the virus strains with mutations”? Please use appropriate terminology.

Response: "...half of the mutated virus strains..." was corrected as “half of the virus strains with mutations”.

17. P. 12, lines 250-252: The authors state that their findings indicate an increase among MSM in Shanghai. How can this conclusion be made? The study that they refer to was performed in Fujian which is situated 800 km from Shanghai. The authors need to rephrase and motivate how this comparison is relevant.

Response: We have rephrased the wordings as follows: Chen’s study revealed that CRF07_BC was increasing among men who have sex with men (MSM) in Fujian; our findings also indicated that CRF07_BC was one of the most prevalent genotypes among MSM in Shanghai.

The reason why we compared our findings with the study that performed in Fujian is just to show the molecular epidemiological characteristics of HIV-1 in different areas across China.

18. The sequences does not appear to have been submitted to Genbank which is highly recommended and common standard – It is also a requirement for publication in BMC Infectious Diseases.
Response: We would very much like to submit the sequences to Genbank, but there are so many sequences that we need more time to do this work. We are very sorry for not finishing the task in time. However, the datasets analyzed during the current study are available from the corresponding author on reasonable request.

19. Table 3: Here the authors report that they have used Fisher's exact test. This was not mentioned in the Methods section. Please clarify.

Response: We have added the description of the statistical method-Fisher's exact test -in the method section. (“…Categorical variables were expressed as frequencies and percentages and compared using the chi-square(x2) test or Fisher exact test….“)

Minor points

1. The manuscript contains several typos and would benefit from some language editing in some sections. It also needs to be reviewed for consistency in terminology and use of symbols and the same font throughout.

Response: Thank you for the detailed comments on the manuscript. We have revised the text accordingly.

2. The section in the introduction that describes the HIV types and groups can be written more concise. For example, the description of HIV-2 and non-M groups on p. 5 is not very relevant for the current study.

Response: In order to describe HIV types and groups more concisely, we deleted the following sentences:

“However, HIV-1 group O, N and HIV-2 cause a small minority of infections in local areas of Africa because of the relatively poor capacity for transmission. Group P ("pending the identification of further human cases") was designated in 2009 when a newly analyzed HIV sequence was reported to have greater similarity to a simian immunodeficiency virus discovered in wild gorillas than to SIVs from chimpanzees.”

3. The term "gender" refers to e.g. each individual's personal identification of one's own gender which is not necessarily the same as the biological sex of that individual. This may seem like semantics, but in scientific writing the use of sex instead of gender is preferred (see e.g. Lancet's instructions for authors).

Response: We used sex instead of gender in the revised manuscript.
4. P. 10, lines 199-200: Revise p-values to p<0.0001.
Response: We have revised p-values to p<0.0001.

5. P. 11, 234-235: This part is redundant from the beginning of the discussions section.
Response: We have removed this part from the discussion section.

6. Table 1: Why is some letters small and some capitalized?
Response: we converted all the lowercase letters into uppercase letters.

Response to Sviatslau Sasinovich, Ph.D. (Reviewer 1):

1. Please revise and proofread the manuscript for the many typographical and grammatical errors;
Response: Thank you for the detailed comments on the manuscript. We have revised the text accordingly.

2. In accordance with BMC infectious diseases guidelines, section "Availability of data and materials" should include information on where data supporting the results reported in the article can be found including, where applicable, hyperlinks to publicly archived datasets analyzed or generated during the study. You indicate only where data can be found but not specify how others can access the data: from the corresponding author on reasonable request? Or if the dataset generated and analyzed during the current study is not publicly available please specify the reason why data is not public available.
See: https://bmcinfectdis.biomedcentral.com/submission-guidelines/preparing-your-manuscript/research-article, go to "Availability of data and materials".
Response: We would very much like to submit the sequences to Genbank, but there are so many sequences that we need more time to do this work. We are very sorry for not finishing the task in time. However, the datasets analyzed during the current study are available from the corresponding author on reasonable request.

3. Methods. It would be better to confirm HIV genotyping with some of the HIV subtyping tools such REGA (http://dbpartners.stanford.edu:8080/RegaSubtyping/stanford-hiv/typingtool/) or COMET HIV-1 (https://comet.lih.lu/index.php);
Response: We have reanalyzed and confirmed HIV-1 genotypes using phylogenetic analysis. HIV-1 pol sequences, together with reference sequences of different subtypes and CRFs, were aligned and further edited manually using the BIOEDIT version 7 (www.mbio.ncsu.edu/BioEdit/bioedit.html). All the subtyping reference sequences were downloaded from the Los Alamos HIV database (https://www.hiv.lanl.gov/content/sequence/HIV/mainpage.html). The phylogenetic trees were generated using the Neighbor-joining method. We added the description of methodology and the results in the revised manuscript accordingly.

4. Conclusions. I would suggest do not use word "extremely" in conjunction with the word "diverse" because there is no clear definition of what is extreme diverse and what is not extreme (or regular) diverse;

Response: The word “diverse” had been removed from the sentence.

5. It is not clear what you mean using words "subjects" and "cases" in your article. For example:

Line 58: Among the 317 successfully amplified cases…

Line 119: A total of 338 subjects were enrolled…

Line 156: Among the 338 enrolled individuals, 317 subjects’ pol gene sequences…

Please, use appropriate words described a subjects and objects of your study: patients, individuals, sequences, etc.

Response: We have revised the text accordingly using appropriate terminology.

6. Lines 46 and 85: Genetic variability and liability to develop drug-resistant mutations are the main characteristics of HIV; The main characteristics of HIV are its enormous genetic variability and rapid evolution… Please, specify, what is the main characteristic of HIV;

Response: We have modified the words as follows:

The main characteristics of HIV are its enormous genetic variability and liability to develop drug-resistant mutations along with high rates of virus replication.

7. Line 51. HIV patients. Better to use "HIV-infected patients";

Response: “HIV patients” were changed to “HIV-infected patients”.
8. Line 125. CD4 cell counts is not related to basic epidemiological data. Moreover, if CD4 cell count was performed from patient's blood sample, please indicate this in Methods. Otherwise specify, how you get CD4 cell number;

Response: We deleted CD4 cell counts from the basic epidemiological data, and indicated the method for CD4 cell count testing.

(“…Basic epidemiological data such as sex, age, marital status, and self-reported transmission route were recorded upon enrollment. CD4+T cell counts were detected by the flow cytometry…”)

9. Line 92. Please provide a reference link to confirm that poor capacity for transmission is the main factor for lower HIV-1 groups N, O and HIV-2 infection numbers.

Response: In order to describe HIV types and groups more concisely, we deleted these words from the introduction section.

10. Lines 219-220 and 234-235. Duplicated information about subtype prevalence in single section;

Response: We have deleted the following words:

“…Of the 12 identified subtypes, CRF01_AE was the most prevalent (50.8%), followed by CRF07_BC (30.3%) and B (8.5%)…”

11. Line 259. Please, describe what policy "Four Free and One Care" means or specify the link where such information can be obtained;

Response: The implication of the policy "Four Free and One Care" is that “free antiretroviral drugs to AIDS patients who are rural residents or people without insurance living in urban areas; free voluntary counselling and testing; free drugs to HIV-infected pregnant women to prevent mother-to-child transmission, and HIV testing of newborn babies; free schooling for AIDS orphans; care and economic assistance to the households of people living with HIV/AIDS).”

12. Line 424, Table 1. Please specify, why some nucleotides (g) in the primer sequences are lowercase. If there is no any reason, please make all primer sequences uppercase;

Response: We have made all the primer sequences uppercase.

13. Line 424, Table 1. Coordinates of MAW 26 (F) primer are 2028-2050, not 2027-2050;
Response: We have revised the numbers to 2028-2050.

14. Line 433, Table 2. Why you provide information only for males in the table?
Response: We added information for female patients.

15. Line 439, Table 3. Please note that table title is 15 words maximum (your is 24). Please modify Table 3 title.
Response: We reedited the table 3 with the title “Multinomial logistic regression analysis for risk factors of pretreatment drug resistance.”

16. Line 439, Table 3 header. "317 newly diagnosed treatment-naive AIDS patients". Do all of the 317 patients have an AIDS? Or probably you mean HIV/AIDS? Please correct.
Response: For the 317 patients, we mean HIV/AIDS patients.

17. Line 439, Table 3. As because the diversity of HIV-1 subtypes is one of the main results of your study it might be better to list in the table all subtypes that you found;
Response: All HIV-1 subtypes found in our study are specified in the manuscript and shown in Figure 2.