PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Neonatal acquisition of extended–spectrum beta-lactamase-producing Enterobacteriaceae in the community of a low-income country (NeoLIC): protocol for a household cohort study in Moramanga, Madagascar
AUTHORS	Harimanana, Aina; Rakotondrasoa, Andriniaina; Rivoarilala, Odile Lalaina; Criscuolo, Alexis; Opatowski, Lulla; Rakotomanana, Elliot Fara Nandrasana; Herindrainy, Perlinot; Collard, Jean-Marc; Crucitti, Tania; Huynh, Bich-Tram

VERSION 1 – REVIEW

REVIEWER	Pokhrel, Bhishma Patan Academy of Health Sciences
REVIEW RETURNED	01-Jun-2022

GENERAL COMMENTS	Please correct the exclusion criteria for neonate (the second exclusion criteria-this should be as -with plans to move away during the period of the study)

REVIEWER	Sands , Kirsty Cardiff University
REVIEW RETURNED	14-Jun-2022

GENERAL COMMENTS	The present a protocol for the multidisciplinary NeoLIC study. This prospective and longitudinal study blends together microbiology, whole genome sequencing, anthropology/ethnographic analyses and statistical and mathematical modelling. NeoLIC will therefore combine detailed microbiology data (quantitative) with rich qualitative and social data. Whilst the NeoLIC protocol has been well presented herein, I have one or two main comments and concerns. Firstly, it would be better to split the strengths and limitations section for clarity. Secondly, the microbiology section is not entirely clear, and I think diagrams/flow charts and figures would help clarify. If over 5,000 samples will be collected in total, the potential microbiology workload is very large. Following enrichment, would what be the expected % growth rate, I would imagine this to be very high, even onto the ESBL chromogenic agar due to the sample type (diverse microbiome, especially adults). Therefore, if each sample presents multiple distinct phenotypic colonies, how many and which will be selected? This is not clear. How will this approach differentiate different strains of the sample species, or different species will

	overlapping chromogens? The limitations section should be expanded. This approach could lead to the phenotypic analysis of several thousand distinct colonies, if between 3-10 are selected from each sample. Thirdly, how will, and how many isolates be selected for whole genome sequencing? The authors should describe this in detail with appropriate justification. Are all potential ESBL isolates to be selected? A particular species/genera? Will the authors employ any molecular screening to determine the presence of certain ESBL markers, i.e. TEM, SHV, OXA-1, CTX-M? Fourthly, the whole genome section does clearly explain how the authors propose to analyse clonal dissemination, but the plasmid analysis may not be sufficient. If certain ARG are carried on plasmids, will the authors explore HGT events and whether plasmids and/or MGE are being disseminated? In order to achieve this, long-read sequencing may need to be employed, and perhaps the limitations could include this, if not possible? Finally, and largely out of interest, the authors could also expand on their reasoning/justification to work in the community/outside of tertiary centres in Madagascar.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Dr. Bhishma Pokhrel, Patan Academy of Health Sciences
Comments to the Author:
Please correct the exclusion criteria for neonate (the second exclusion criteria-this should be as -with plans to move away during the period of the study)
We thank you for noticing this mistake which has been corrected.

Reviewer: 2
Dr. Kirsty Sands , Cardiff University
Comments to the Author:
The present a protocol for the multidisciplinary NeoLIC study. This prospective and longitudinal study blends together microbiology, whole genome sequencing, anthropology/ethnographic analyses and statistical and mathematical modelling. NeoLIC will therefore combine detailed microbiology data (quantitative) with rich qualitative and social data.

Whilst the NeoLIC protocol has been well presented herein, I have one or two main comments and concerns.

Firstly, it would be better to split the strengths and limitations section for clarity.
The section on strengths and limitations has been split.

Secondly, the microbiology section is not entirely clear, and I think diagrams/flow charts and figures would help clarify. If over 5,000 samples will be collected in total, the potential microbiology workload is very large. Following enrichment, would what be the expected % growth rate, I would imagine this to be very high, even onto the ESBL chromogenic agar due to the sample type (diverse microbiome, especially adults). Therefore, if each sample presents multiple distinct phenotypic colonies, how many

and which will be selected? This is not clear. How will this approach differentiate different strains of the sample species, or different species will overlapping chromogens? The limitations section should be expanded. This approach could lead to the phenotypic analysis of several thousand distinct colonies, if between 3-10 are selected from each sample.
Based on our previous experiences, we estimated an ESBL growth rate of 30%. Thus, we expected to isolate overall a maximum of 1500 ESBLs. Indeed, the workload was large and human resources, both in the field and in the microbiology laboratory, have been recruited accordingly. Also, the number of households to be included at a monthly basis (4 households per month) and the duration of the inclusion period (15 months) were determined so as not to overload the teams.

We added the following sentence in the sample size section to make this clear:

Based on our previous experiences, we estimated an ESBL growth of 30%. Thus, among the 5040 estimated samples to be collected, we expect to isolate a maximum of 1512 ESBLs. The human resources required, the number of households to be included at a monthly basis (4 households per month) and the duration of the inclusion period (15 months) were determined so as not to overload both the teams working in the field and in the laboratory.

We also added in the bacteriology section the following sentence regarding the selection of the colonies:

The colonies will be selected based on their color and morphological appearance such as size, surface and edge. Based on our previous experiences we expect that in stool the majority of the colonies will belong to the Enterobacteriaceae and will be pink (Escherichia coli) or metallic blue (Klebsiella spp) and that mixed cultures containing 2 or more morphologically distinct colonies will be frequent. Colonies from each color present will be selected with the exception of blue colonies of which a maximum of three colonies will be selected each showing a different appearance and with a preference for the metallic blue colonies, if applicable. Cultures showing mixtures of colonies will be sub cultured until pure isolated colonies are obtained for further processing. Only single and pure bacterial colonies will be processed for identification and if Enterobacteriaceae also for antibiotic susceptibility testing and DNA extraction for WGS.

Thirdly, how will, and how many isolates be selected for whole genome sequencing? The authors should describe this in detail with appropriate justification. Are all potential ESBL isolates to be selected? A particular species/genera? Will the authors employ any molecular screening to determine the presence of certain ESBL markers, i.e. TEM, SHV, OXA-1, CTX-M?
All potential ESBL isolates belonging to the Enterobacteriaceae will be selected for WGS. We will not perform molecular screening.
The following sentence has been added:

All ESBL belonging to the Enterobacteriaceae isolates will be selected for WGS.

Fourthly, the whole genome section does clearly explain how the authors propose to analyse clonal dissemination, but the plasmid analysis may not be sufficient. If certain ARG are carried on plasmids, will the authors explore HGT events and whether plasmids and/or MGE are being disseminated? In order to achieve this, long-read sequencing may need to be employed, and perhaps the limitations could include this, if not possible?

We agree with the reviewer, we will not be able to conduct “plasmid tracking” per se as long reads will not be obtained.

However, we will conduct scaffold classification using Platon to determine if each resistance gene was carried on the chromosome or plasmid (s).
We added the following sentences in the methods and limitations sections, respectively.
Thanks to the scaffold classification using Platon, each determined resistance gene will be assessed as carried on the chromosome or plasmid(s), therefore allowing the tracking of possible resistance gene transfers from plasmid to chromosome.
We will not be able to conduct plasmid tracking per se as long-read sequencing will not be performed.

Finally, and largely out of interest, the authors could also expand on their reasoning/justification to work in the community/outside of tertiary centres in Madagascar.

We added the following sentence in the introduction
Data from the community are scarce and available ones show high ESBL-PE carriage prevalence. In Madagascar, the estimated prevalence of ESBL-PE carriage in pregnant women was 18.5% (95% Confidence Interval (CI), 14.5% -22.6%) (compared with <5% in Europe). These findings demonstrate the need to work at the community level where the ESBL-PE prevalence is significant.

VERSION 2 – REVIEW

REVIEWER	Sands , Kirsty Cardiff University
REVIEW RETURNED	06-Aug-2022

GENERAL COMMENTS	Thank you. All comments have been appropriately addressed in the resubmission.