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)	Cohort profile: The Mâncio Lima cohort study of urban malaria in Amazonian Brazil
AUTHORS	Johansen, Igor; Rodrigues, Priscila; Tonini, Juliana; Vinetz, Joseph; Castro, Marcia; Ferreira, Marcelo

VERSION 1 – REVIEW

REVIEWER	Dinka, Hunduma Adama Science and Technology University, Applied Biology
REVIEW RETURNED	23-Mar-2021

GENERAL COMMENTS	The content and the findings are not as per it's title. And even the findings are not novel.

REVIEWER	Tshefu Kitoto, Antoinette University of Kinshasa
REVIEW RETURNED	09-Apr-2021

GENERAL COMMENTS	Comparing Tables 2 and 3 some periods of field work do not match : Table 2 , Mar-Apr 18 for visit 1 whereas it is April-May on Table 3 In the Findings To Date Section on page 12, at the last sentence you say that your PCR protocol detects up to 10 times more malaria infection than microscopy in cohort participants but the results are not presented

REVIEWER	Martins , MDRO Universidade Nova de Lisboa, Instituto de Higiene e Medicina Tropical
REVIEW RETURNED	03-Aug-2021

GENERAL COMMENTS	INTRODUCTION The rationale for the study is well described; there are no specific research questions but the objectives are clearly stated and we understand the motivation of the study. The authors mentioned (in the last sentence) that the study has been expanded to include SARS- CoV-2 antibody measurements, however there are no results on this specific topic. COHORT DESCRIPTION The setting, locations and relevant dates are described in detail, including periods of recruitment, exposure and follow-up. However, under the section "Study site" (line 38-50) the authors described the funding and also the context in which the study has

	been developed. In my opinion this paragraph should be moved from here to a subsection Funding. The eligibility criteria and recruitment is fine. A flow diagram is presented and is clear. Reasons for non-participation are somehow reported and are mainly related to people that migrate. Other reasons are not stated. Methods of data collection are stated; however, because there are two level of information (at the individual level and at the household level) it should be clear what data is collected at the household level. Methods of follow-up are well described. The authors referred (line 5 page 10) that a questionnaire is applied to obtain and update demographic, socioeconomic, occupational, behavioral, and morbidity information and measures are listed in Table 1. I think that is not clear in the questionnaire variables, those who are collected at the household level, adult level, children level, etc.. This table must be improved with the content of the questionnaire at the different levels and it also should be mentioned the number of participants with missing data for each variable of interest. FINDINGS TO DATE Some of the characteristics of study participants are presented; mainly demographic, geographical and clinical; however, findings related to socioeconomic, occupational and behaviour information are missing. This is a very important information that must be added in Table 3 or in another Table.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
The content and the findings are not as per it's title. And even the findings are not novel.

We have formatted the manuscript according to the BMJ Open□s instructions for cohort profile articles. Indeed, the paper includes a few published results in the section “Findings to Date”, but the overall description of the population-based cohort has been submitted for the first time.

Reviewer: 2
Comparing Tables 2 and 3 some periods of field work do not match : Table 2 , Mar-Apr 18 for visit 1 whereas it is April-May on Table 3

Many thanks for pointing this out. In fact, information given in the footnote of Table 3 is correct. We corrected Table 2 and added information about dengue serology.

In the Findings To Date Section on page 12, at the last sentence you say that your PCR protocol detects up to 10 times more malaria infection than microscopy in cohort participants but the results are not presented.

Yes, the reviewer is correct. We decided not to present preliminary PCR results in this paper because we are currently validating our assays. We screened all samples for malaria parasites with a genus-based diagnostic PCR and are now applying a newly developed species-specific protocol to measure the prevalence of P. falciparum and P. vivax infections. We do not consider these results as definitive. We changed the main text to:

“Ongoing analyses indicate that our PCR protocol detects up to 10 times more malaria infections than microscopy in cohort participants, but further standardization and validation are in progress.”

Reviewer: 3

INTRODUCTION
The rationale for the study is well described; there are no specific research questions but the objectives are clearly stated and we understand the motivation of the study.
The authors mentioned (in the last sentence) that the study has been expanded to include SARS-CoV-2 antibody measurements, however there are no results on this specific topic.

We added the following information regarding COVID-19 and its relationship with prior dengue exposure in our cohort:

"We have recently shown that serologically proven prior dengue infection is associated with increased subsequent risk of clinically apparent COVID-19 in this cohort.17 Dengue IgG antibodies were detected in 37.0% of the 1,285 cohort participants tested in October-November, 2019, with 10.4 seroconversion events per 100 person-years over the following 12 months. In October-November, 2020, 35.2% of the participants tested had anti-SARS-CoV-2 IgG and 57.1% of the 448 SARS-CoV-2 seropositives reported clinical manifestations of COVID-19 at the time of infection. Participants aged >60 years were twice more likely to have symptomatic COVID-19 than under-five children. Importantly, prior dengue infection was associated with twice the risk of clinically apparent COVID-19 upon SARS-CoV-2 infection after adjustment for identified confounders.17"

Note that the revised version of the introduction (last paragraph) now reads:

"The original study has since expanded to include SARS-CoV-2 antibody measurements during the ongoing COVID-19 pandemic in this hard-hit region and investigate possible interactions between dengue and COVID-19.17

The revised version of Table 1 now mentions dengue serology carried out in the visits 4 and 5.

COHORT DESCRIPTION

The setting, locations and relevant dates are described in detail, including periods of recruitment, exposure and follow-up.
However, under the section "Study site" (line 38-50) the authors described the funding and also the context in which the study has been developed. In my opinion this paragraph should be moved from here to a subsection Funding.

We moved this information to the Funding section:

"The Mâncio Lima cohort study is part of the National Institutes of Health (NIH)-funded Amazonian International Center of Excellence for Malaria Research network (https://www.niaid.nih.gov/research/amazonian-international-center-excellence-malaria-research) funded by the National Institutes of Health (NIH), United States (Amazonian International Center of Excellence for Malaria Research, U19 AI089681 to J.M.V.) (…)".

The first sentence of the Study Site section now reads:

"The Mâncio Lima cohort study aims to investigate malaria epidemiology, vector biology and ecology, diagnostics, transmission biology, and clinical pathogenesis in Amazonian Brazil."

The eligibility criteria and recruitment is fine. A flow diagram is presented and is clear. Reasons for non-participation are somehow reported and are mainly related to people that migrate. Other reasons are not stated.

We added to the description of follow-up: “Participants who moved away from the study site and those who withdrew their consent to participate were lost for follow-up.”

Methods of data collection are stated; however, because there are two level of information (at the individual level and at the household level) it should be clear what data is collected at the household level. Methods of follow-up are well described.

We have reordered the text to make it clearer:

“Structured questionnaires have been applied to study participants during study visits to obtain and update the demographic, socioeconomic, occupational, behavioral, and morbidity information listed in Table 1. Dates of follow-up visits and the number of participants interviewed in each visit are shown in Table 2. Both individual and household-level information was collected during study visits. GPS coordinates were obtained for all dwellings. Data were entered using tablets programmed with REDCap25 and subsequently exported to Stata SE 15.0 (StataCorp, College Station, USA) for statistical analysis. Because study participants are nested into households, which introduces dependency among observations, we have been using mixed-effects logistic or Poisson regression models with random effects at the household level and robust variance for data analysis.”

The authors referred (line 5 page 10) that a questionnaire is applied to obtain and update demographic, socioeconomic, occupational, behavioral, and morbidity information and measures are listed in Table 1. I think that is not clear in the questionnaire variables, those who are collected at the household level, adult level, children level, etc.. This table must be improved with the content of the questionnaire at the different levels and it also should be mentioned the number of participants with missing data for each variable of interest.

We now indicate in Table 1 which key variables are individual (one value for each participant) or household-level (the same value attributed to all household members). We do not mention the number of participants with missing information for each variable because this would make Table 1 very confusing, but added to the main text the following: “For the vast majority of variables, information is missing for <5% of participants.” Moreover, the number of participants with missing information for selected variables is shown in the new Supplemental Table 2.

FINDINGS TO DATE

Some of the characteristics of study participants are presented; mainly demographic, geographical and clinical; however, findings related to socioeconomic, occupational and behaviour information are missing. This is a very important information that must be added in Table 3 or in another Table.

We added this information to a new table (Supplemental Table 2 in the revised version).

VERSION 2 – REVIEW

REVIEWER	Martins , MDRO Universidade Nova de Lisboa, Instituto de Higiene e Medicina Tropical
REVIEW RETURNED	10-Sep-2021

GENERAL COMMENTS	The authors answer to all my questions in a correct way. I think the paper can be accepted