Assessing the short, intermediate and long-term health effects of COVID-19 on the survivors in Zambia: a prospective quantitative study protocol

Cephas Sialubanje, Evaristo Kunka, Phyllis Ingutu Sumbwa, Nawa Mukumbuta, Margaret Chibowa, Gardner Syakantu, Laston Chikoya, Victor M Mukonka

ABSTRACT

Introduction Little is known about the clinical course of COVID-19 following mild symptoms, and how the disease affects the survivors over time. Moreover, information on the severity of the long-term health effects as well as the associated risk factors is scant. This study aims to determine the short, intermediate and long-term health effects of COVID-19 on the survivors and the associated risk factors.

Methods and analysis We propose conducting a 24-month prospective quantitative study in 10 health facilities (2 specialist, 3 regional, 2 mission and 3 subdistrict hospitals) from Lusaka and Southern Province of Zambia. Health facilities will be those which served COVID-19 treatment centres during the third wave (June–August 2021). Study participants will comprise a randomly selected cohort of 450 COVID-19 survivors who had mild or no symptoms (80%) and severe cases (20%). Using a questionnaire, respondent demographic, clinical and laboratory data will be collected at baseline and at a 3-month interval for 18 months using a questionnaire. Respondents’ medical records will be reviewed and data collected using a checklist. Descriptive statistics will be computed to summarise respondents’ characteristics and clinical outcomes. Bivariate analysis (X² and t-test) will be conducted to test the association between respondent characteristics and clinical outcomes. Multivariate logistic regression analysis will be run to determine the risk factors for short, intermediate and long-term health effects; adjusted ORs will be computed to test the strength of the association (p<0.05).

Ethics and dissemination Ethical approval was obtained from the University of Zambia Biomedical Research and the National Health Research Authority. Results will be disseminated to key stakeholders in Zambia, international conferences, and likely lead to design of evidence-informed strategies to mitigate health effects of COVID-19 on survivors.

INTRODUCTION

Globally, as of 23 August 2022, there have been 594 367 247 confirmed cases of COVID-19 and 6 451 016 deaths. In Zambia, a total of 325 248 confirmed cases and 4003 deaths have been reported. Although higher case fatality rates were reported from developed countries during successive waves, most infections in Zambia were mild or an overall case fatality rate of 1.23%, except during the third wave (June–August 2021) when a higher case fatality rate, attributed to the delta variant, was reported. Many cases presented with no or mild symptoms including fever, dry cough and tiredness. Other less common symptoms were aches and pains, sore throat, diarrhoea, conjunctivitis, headache, loss of taste or smell, skin rash, or discolouration of fingers or toes.

Although most people typically recover from COVID-19 after 2–6 weeks, studies have demonstrated that some symptoms may recur...
or persist for weeks or months following initial recovery among both the healthy people and those with mild disease. Some medical complications may have lasting health effects on the survivors; some patients may not regain their full health. For example, evidence from the WHO indicates that approximately 10%–15% of cases progress to severe disease; 5% become critically ill with severe clinical manifestations and complications ranging from shortness of breath and general breathing difficulties to organ failure and/or death. Respiratory and cardiac complications have been reported to be the most common complications and leading causes of death, especially among the elderly and those with comorbidities. Although elderly patients and those with comorbidities have a high probability of developing complications, a Centers for Disease Control and prevention (CDC) report indicates cases of healthy adults and young people who developed COVID-19-related complications. Nevertheless, due to health system challenges in low/middle-income countries, most disease complications and long-term effects among survivors with mild or no symptoms may go unreported or undiagnosed.

Thus, much is still unknown about the clinical course of COVID-19 following a mild or asymptomatic infection with regard to: (a) which patients will have some symptoms persist or recur; (b) long-term effects of the disease and (c) the associated risk factors. Current research has focused on understanding the biology of the virus itself, prevention and control measures, clinical manifestations of the disease and other related aspects; little attention has been paid to the long-term effects of the disease. In view of these important knowledge gaps on the clinical course of the disease among the individuals with mild or no symptoms, a prospective study is needed to follow up a selected cohort of COVID-19 survivors who had severe, mild or no symptoms. The proposed study will attempt to answer the following questions: (a) What are the long-term health effects of COVID-19 on the survivors?; (b) What is the severity of the long-term health effects of COVID-19 on the survivors?; (c) What are the risk factors for developing long-term health effects of COVID-19 on the survivors?

The aim of this study is to investigate the short, intermediate and long-term health effects of COVID-19 on the survivors in order to inform policy and development of clinical management, counselling and health promotion messages for improved long-term care of COVID-19 survivors in Zambia.

The specific objectives are to:
1. Identify the short, intermediate and long-term health effects of COVID-19 on survivors.
2. Describe the severity of the health effects of COVID-19 on the survivors.
3. Determine the risk factors for the short, intermediate and long-term health effects of COVID-19 on the survivors.

If successful, the study findings will bridge the knowledge gaps and inform design of policy and clinical guidelines for the management and care of COVID-19 survivors. Currently, information on the short, intermediate and long-term effects of COVID-19 and their severity on the survivors as well as the risk factors associated with these effects is scant.

**METHODS AND ANALYSIS**

**Study designs**

We propose conducting a 24-month prospective study (September 2021–August 2023), employing quantitative (survey) methods and medical records review as data collection techniques. Corroborating two data collection techniques is likely to increase the validity of the study findings.

**Study setting**

The study will be conducted in 10 health facilities (2 specialist, 3 regional, 2 mission and 3 subdistrict hospitals) from Lusaka and Southern Province of Zambia. To be selected, health facilities will be those which served as COVID-19 treatment centres during the third wave (June–August 2021). The 10 health facilities will be selected from a total of six districts: three from Lusaka (Kafue, Chirundu and Lusaka districts) and three from Southern Province (Choma, Monze and Livingstone districts) (see table 1).

Lusaka has seven districts, including Lusaka city, the commercial capital and government headquarters, with a total population of 3.19 million; Southern Province has a total population of 1.85 million. Both provinces provide specialised, second and primary-level healthcare services. Lusaka has two specialised, five second-level and several primary-level healthcare facilities. Southern Province has one specialist hospital, six level two, several district hospitals and primary health facilities. In addition, there are private and faith-based healthcare providers in each province. The Ministry of Health (MoH) head-quarters and the Zambia National Public Health Institute

<table>
<thead>
<tr>
<th>Table 1 Study sites</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Name of institution</strong></td>
</tr>
<tr>
<td>Specialist hospital</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Regional hospital</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Mission hospital</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Subdistrict hospital</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Participants and sampling techniques

Study participants will be selected from people who tested positive to and survived COVID-19 during the third wave (June–August 2022). First, a complete list of people who tested positive for COVID-19 during the third wave at each study site will be obtained from the list line which will serve as a sampling frame. The third wave has been selected based on the large number of people who got infected; it will also be chosen for logistical purposes (that is, to make the study feasible). Next, a stratified random sampling technique will be used to select respondents who tested positive for COVID-19. This approach will ensure that both survivors who had mild or no symptoms and those who had severe disease are selected. Stratification will be done in the ratio 1:4 (mild or no symptoms 80%; severe disease 20%) according to the WHO guidelines. Selected participants will then be contacted on phone and email to inform them about the study and confirm their willingness to be recruited. Those who agree to be included in the study will be followed to their homes or any place of their choice for enlisting into the study. To be included into the study, participants will need to:

- Have tested positive for COVID-19 during the third wave (June–July 2021).
- Be 15 years and above of age.
- Have been residing in the selected study area for not less than 3 months.
- Those with active COVID-19 infection will not be eligible to participate.

Sample size estimation

A single-population proportion formula was used to estimate the sample size required for the study. The formula uses descriptive studies and takes into consideration the following parameters: prevalence of the condition of interest, desired non-response rate and the margin of error. Since the actual prevalence of cases with severe and mild disease is not known, an assumption of 50% will be made with a 5% margin of error and 10% non-response rate as shown below:

\[ n = \left( \frac{Z^2 \times (P(1-P))}{d^2} \right) \]

Where, \( n \) = sample size;
\( Z \) = z-score at 95% confidence level (1.96);
\( P \) = proportion of population with a characteristic of interest (50%);
\( d \) = margin of error (at 5%).

Allowing for a 10% non-response rate, a total sample size of 423 respondents (approximately 450) will be required in our study (90 with severe disease and 360 with mild or no symptoms) as shown in Table 2 below.

Table 2 Sample size per study site

<table>
<thead>
<tr>
<th>District</th>
<th>Mild or not symptomatic</th>
<th>Severely symptomatic</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lusaka (Matero)</td>
<td>36</td>
<td>9</td>
<td>45</td>
</tr>
<tr>
<td>Lusaka (Chelstone)</td>
<td>36</td>
<td>9</td>
<td>45</td>
</tr>
<tr>
<td>Lusaka (Chilenje)</td>
<td>36</td>
<td>9</td>
<td>45</td>
</tr>
<tr>
<td>Levy Mwanawasa University Teaching Hospital (LMUTH)</td>
<td>36</td>
<td>9</td>
<td>45</td>
</tr>
<tr>
<td>Mazabuka</td>
<td>36</td>
<td>9</td>
<td>45</td>
</tr>
<tr>
<td>Kafue</td>
<td>36</td>
<td>9</td>
<td>45</td>
</tr>
<tr>
<td>Chirundu</td>
<td>36</td>
<td>9</td>
<td>45</td>
</tr>
<tr>
<td>Monze</td>
<td>36</td>
<td>9</td>
<td>45</td>
</tr>
<tr>
<td>Choma</td>
<td>36</td>
<td>9</td>
<td>45</td>
</tr>
<tr>
<td>Livingstone</td>
<td>36</td>
<td>9</td>
<td>45</td>
</tr>
<tr>
<td>Total</td>
<td>360</td>
<td>90</td>
<td>450</td>
</tr>
</tbody>
</table>

Variables

The study variables will include the outcome and independent variables. The outcome variables will be the effect of COVID-19 on the survivor and severity of complications. Effects of COVID-19 on survivors will be classified as short term, intermediate or long term based on the time of manifestation following infection (within 3 months, 6 months, 12 months or more, respectively), according to the WHO guidelines. Severity of complications will be classified as mild, moderate and severe, based on the clinical and radiological/laboratory staging. Independent variables will be: (a) demographic information: age, sex, number of children, place of residence; (b) socioeconomic information: level of education, occupation, level of income; (c) clinical symptoms; (d) laboratory tests/results; (e) type of complications (actual diagnosis), time complication diagnosed.

Confounding variables will be: (a) vaccination status, (b) pre-existing medical conditions and (c) COVID-19 reinfection.

Data sources and collection procedures

Data will be collected from study respondents and patient clinical records. First, demographic and clinical data will be collected at baseline. Next, clinical data will be collected at a 3-month interval (that is, at baseline, months 3, 6, 9, 12, 15 and 18), under the supervision of the principal investigator (PI) and two field supervisors. A data collection instrument (online supplemental material 1), adapted from the NCL Primary Care COVID-19 Recovery Questionnaire,17 Derbyshire Post COVID-19.
Syndrome Questionnaire and Post COVID-19 Recovery Clinic Baseline Questionnaire will be used. To ensure internal validity, the data collection tool will be translated into the local languages, Bemba, Nyanja and Tonga. Activities for the first visit will include provision of detailed information about the study, obtaining informed consent and collection of respondent baseline data, including: (a) demographic information, (b) COVID-19 diagnosis, (c) vaccination status, (d) medical history, (e) current symptoms, if any, (f) severity of symptoms, (g) hospital admission for current symptoms, (h) laboratory and radiological investigations conducted and results, (i) medical diagnosis if known, (j) medical treatment for current symptoms. To be objective, baseline clinical and laboratory information will be collected from respondents and medical records from the health facilities. Activities for subsequent follow-up visits will include data collection on: (a) respondents’ general well-being, (b) existing and new symptoms, (c) diagnosis of complication, (d) clinical manifestations/symptoms, (e) time of onset of symptoms from a positive COVID-19 result, (f) severity of complications, (g) laboratory/pathology investigations, (h) radiological investigations conducted including chest X-ray and echocardiography, and (i) overall survivor outcomes.

**Hiring and training of data collectors**
A total of eight research assistants (four male and four female) will be trained. Research assistants will be hired from the Schools of Public Health at University of Zambia and Levy Mwanawasa Medical University. Hiring and training of data collectors will be done by the PI and members of the research team after ethical approval, before commencement of the actual data collection process. To be cost-effective, training will be conducted in a central place, preferably Lusaka city. The training will take a total of 5 days and will be in two phases. Phase one will be theory in research and take 3 days. The main topics will include: (1) basic principles of research, (2) purpose and objectives of the current study, (3) data collection and interviewing techniques, (4) informed consent in research involving human subjects and (5) data collection instruments (survey questionnaire and checklist). Phase two will be a field practical and will include pretesting data collection tools and consent form, and will take 2 days. After the field practical, instruments will be revised based on the collected and analysed data and feedback from the research assistants.

**Biases**
Possible biases in the study include selection bias, measurement error and confounding due to vaccination and COVID-19 reinfection. To minimise selection bias, systematic stratified sampling will be used to select participants from the line lists maintained in the study sites. In addition, a group of research team members will work together to select the study sample. To minimise measurement error, research assistants will be trained in data collection techniques. Research assistants will work under the supervision of the field supervisors. A pretested data collection instrument, translated into the local language, will be used. Confounding due to vaccination status, pre-existing medical conditions and COVID-19 reinfection will be controlled for during data analysis.

**Statistical methods**
Data from the survey and medical records review will be entered into an Excel sheet and saved on a password-protected computer. After cleaning up, data will be transferred into SPSS V.21 (IBM) for analysis. Descriptive statistics will be used to summarise sociodemographic and clinical data. Bivariate analysis (X^2) will be conducted to test the association between respondent characteristics and clinical outcomes. Multivariate logistic regression analysis will be conducted to determine the risk factors for short, intermediate and long-term health effects as well as their severity; confounding due to pre-existing medical conditions, vaccination status and COVID-19 reinfection will be controlled for and adjusted ORs will be computed to test the strength of the association (p<0.05).

**Patient and public involvement**
The study design was determined by the call for research proposals on COVID-19 research under the strategic research fund. Thus, participants and the public were not directly involved in the conceptualisation and design of the study. However, selection of the study sites was done in consultation with the MoH and ZNPHI. In addition, selection of study participants was done in collaboration with the provincial and district health managers. First, an inception meeting was held with the ZNPHI and MoH headquarters staff to discuss the design and objectives of the study. Next, prefield meetings were held with the two provincial health office staff members to select the study sites. Recruitment of participants was done in conjunction with site managers.

**Ethics and dissemination**
Ethical clearance and approval was obtained from the University of Zambia Biomedical Research Ethics Committee (UNZA BREC 2565-2022). Permission to conduct the study was granted by the National Health Research Authority. Before being recruited into the study, the purpose, process and duration of the study, benefits and associated risks, privacy and confidentiality will be explained to the respondents. Data collectors will explain that there are no direct individual benefits to participating in the study. They will also explain potential risks the study may pose to study participants. Participation in the study may cause some discomfort from answering certain questions, particularly if they had an adverse COVID-19 health outcome. To reduce the risk of disclosure of personal or sensitive information, data collectors will be trained to ensure that participants are helped not to share anything that they are not comfortable with, and that they do not have to respond to any question unless they feel comfortable doing so. Participants will also be...
told that they are free to stop the discussion at any time if they need to. Moreover, participation may expose the study participants and study team to COVID-19 infection. To minimise the risk of infection to the research team and study participants, bio-safety measures will be taken during data collection including social distancing and conducting virtual interviews. In addition, data collectors and study participants will be encouraged to consistently use face masks and hand sanitisers before and after each interview. Used face masks will be disposed of according to the health guidelines. Data collectors will also be trained to minimise to the greatest extent possible any potential discomfort or harm to the participants during all study activities. The study team will minimise any waiting by participants by scheduling appointments during times convenient to participants and ensuring that interviews are kept to as short a time as possible. Those who accept to participate will be asked to sign the consent form (online supplemental material 2) and be enlisted in the study. Written informed consent will be obtained from each respondent. Those who cannot read or write will be asked to mark with an ‘X’. The consent form will be translated into the local language. Participants aged below 18 years will require both parental consent and assent. Privacy of participants and confidentiality of collected data will be assured throughout the study. Interviews will be carried out in participants’ private homes or in the participant’s place of choice. Data collection will not be conducted until we confirm that the location is acceptable to the participants. To ensure data quality, weekly research meetings will be held to ensure effective and efficient execution of the project. On completion of each data collection phase, all files will be stored on a secure passcode-protected computer kept by the PI. At the end of the data collection process, data will be de-identified and the linking file with identifiable data and basic demographics will be stored in a separate file. Only the PI and some identified research team members will have access to the identifiable data. All analyses will be conducted on de-identified data.

If the study is successful, the findings will be disseminated among the stakeholders. A report will be written at the end of the project and shared with key stakeholders, including the funding organisation and MoH. In addition, study findings will be disseminated to key stakeholders in Zambia, then through open-access peer-reviewed journals, websites and international conferences.

Limitations
This study has several limitations. First, the study will only include 450 participants from 10 sites in two provinces, Southern Province and Lusaka (five sites from each province). The study findings may not be generalisable to other patients with different clinical or epidemiological characteristics. To make the study findings more representative, a large-scale study is needed. However, the purpose of the current study is to determine the short, intermediate and long-term health effects and severity of COVID-19 on the survivors and the associated risk factors. Second, although study participants will be sampled from those who had COVID-19 during the third wave (June–August 2021), the follow-up period will run through the fourth wave of COVID-19 (November 2021–February 2022). Some participants may have been infected during the fourth wave as well; it is not clear how reinfection may affect participant health outcomes. Third, the country has been conducting a mass national COVID-19 vaccination campaign since April 2021. It is not clear how participants’ pre-morbid vaccination status affects the health outcomes for the COVID-19 survivors. Furthermore, being prospective in design, the study will follow up study participants for 18 months and some participants may be lost to follow-up. In addition, it is currently not known how long some health effects may take to manifest; some patients may remain asymptomatic within the follow-up period.

Nevertheless, we believe that the study is likely going to contribute to the scientific body of knowledge on the epidemiological and clinical course of COVID-19 following an asymptomatic, mild or severe disease. The findings are also likely to provide in-depth information on the short, medium and long-term health effects of the disease on the survivors, and the associated risk factors and ultimately lead to design of evidence-informed strategies to improve the long-term health effects of COVID-19. Currently, no such study has been conducted in the country and information on this subject is scant.

Contributors CS led the design of the study, drafting of the study protocol and implementation of the study, and drafted this manuscript. EK, PIS and NM contributed to the development of the data collection instruments, enlisting the study sample, and coordinated data collection process. MC, GS, LC and VMM contributed to the revision of the manuscript. All authors read and approved the manuscript.

Funding This work was supported by the National Science and Technology Council (grant number: NSTC 101/6/8), as part of the Emergency COVID-19 research grant under the strategic research fund.

Disclaimer The funder had no influence on the design and implementation of the study, nor on the draft of this manuscript.

Competing interests None declared.

Patient and public involvement Patients and/or the public were involved in the design, conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Obtained.

Provenance and peer review Not commissioned; externally peer reviewed.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is

properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

Author note Costs and payments: There is no cash payment provided to participants for any portion of the study. Participants volunteer the time taken to complete this survey. As a small token of appreciation for their time and opportunity costs, participants will receive a bottle of water or soft drink, in line with UNZA BREC procedures.

ORCID ID
Cephas Sialubanje http://orcid.org/0000-0002-9077-1436

REFERENCES
4 Day M. Covid-19: four fifths of cases are asymptomatic, China figures indicate. BMJ 2020;369:m1375.