Ebola Data Platform - Data Access Application Form

Please review the Data Access Guidelines and the Data Transfer Agreement* before completing this form. Note that the details of all approved applications will be made publicly available on the Ebola Data Platform website.

Please complete all sections of this form *fully* and return to ebolaDAC@iddo.org with the following attachments:

- Academic CV of the Lead Requestor (any format)
- Conflict of Interest Forms completed by the Lead Requestor and each of the Co-applicants listed

SECTION A: RESEARCHER / RESEARCH TEAM INFORMATION	
Lead Requestor Details *(please attach an academic CV)*	
Title	Mr.
First name (given name)	James Jr
Surname (family name)	Njong
Gender	Male
Position at employing organisation/ institution	Research Assistant
ORCID ID https://orcid.org/	
Email	James.berinyuy@aims-cameroon.org
Telephone/Skype/WhatsApp	+237 674 811 226
Employing Organisation/Institution *Institution with a remit including health, research or academic pursuit, and with legal status which includes the scope to sign the **Data Transfer Agreement****	
Institution name	African Institute for Mathematical Sciences
Address	Crystal Gardens, Limbe. South West Cameroon. P.O Box 608 Limbe
Department (if applicable)	NA
Please acknowledge that your institution agrees to execute the Data Transfer Agreement (in the case of your application being approved)	YES (delete as appropriate)

* The ***Data Transfer Agreement*** *is a contract between the University of Oxford (on behalf of IDDO) and the recipient institution that governs the legal obligations and restrictions, as well as compliance with applicable laws and regulations, related to the **transfer** of such **data** between the parties. The named Institution will be required to sign the data transfer agreement before the release of any data by IDDO.*

Co-applicants (ALL individuals accessing the data must be listed. Any additions must be notified to the Ebola DAC) *Add rows as necessary.* *Please attach copies of the Conflict of Interest Form, completed by each of the individuals above.*	
1. Name	Christabel Lemukong Ngufor
1. Title	Student
1. Organisation/Institution	African Institute for Mathematical Sciences
2. Name	Mavis Amoa-Dadzeasah
2. Title	Student
2. Organisation/Institution	African Institute for Mathematical Sciences
3. Name	Christiana Kartsonaki
3. Title	Statistician
3. Organisation/Institution	University of Oxford
4. Name	Trokon O. Yeabah
4. Title	Researcher
4. Organisation/Institution	National Public Health Institute of Liberia
5. Name	Mahamoud Sama Cherif
5. Title	Researcher
5. Organisation/Institution	Gamal Abdel Nasser University of Conakry, Guinea
SECTION B: RESEARCH PLAN	
Title of Proposed Research	Factors associated with mortality in patients Ebola virus disease
Is this a re-submission of a previous application that has been reviewed by the Ebola DAC? If so, please provide the surname of the Lead Requestor and submission date of the previous application.	Yes, it was provisionally approved by the DAC April 9, 2020, pending confirmation of grant funding to Dr Cherif, the lead applicant. As this funding was not awarded, this re-submission is tailored to currently available funding as a part of a capacity building initiative between AIMS and the University of Oxford. Dr Cherif will remain a collaborator, and the analysis will be delivered by AIMS Master of Statistics students, supervised by Oxford Senior staff.

Summary of Research in Lay Language *(suggested ~ 100 words)*

Sparse data available from extremely challenging conditions during the 2014-2016 epidemic has shown several factors are predictive of mortality. These include: high viral load, age of the patients (infants, young children <5 years age and elderly > 60 years), haemorrhagic fever, interval between onset of symptom and admission (delayed admission), pregnancy, symptom of confusion, chest pain, coma, and co-infection with malaria.

In an emergency outbreak situation, it is important to have a reliable indicator for medical screening of the patients and monitoring the evolution of the signs and symptoms

associated with the disease to identify those at substantially increased risk of death. This can ensure optimisation of palliative and curative care delivery.

Scientific Summary of Research *(suggested maximum 300 words)*

This application seeks to characterise the evolution of the signs and symptoms of EVD using the individual patient data hosted by the IDDO Ebola data platform. Individual participant data meta-analysis (IPD-MA) is currently considered to be the gold-standard approach for synthesis of results across several studies and provides a unique opportunity to characterise the evolution of the disease symptoms.

Outcome will be each of the following symptoms or signs: fever, bleeding, confusion, conjunctivitis, intense fatigue, hiccups, vomiting, diarrhoea, anorexia, hypovolemic shock, anaemia, temperature, blood pressure; and clinical biomarkers: white blood cell count, haemoglobin, platelet, C-reactive protein, electrolytes, creatinine, liver transaminases, CPK, lactic acid, base excess, PTT; viral load, PCR cycle threshold antibody titres.

Summary of Research Objectives *(suggested maximum 200 words)*

Primary objective

- To describe clinical signs and symptoms at baseline and during hospitalisation period in ETC in patients with confirmed EVD

Secondary objectives

- To carry out longitudinal mixed effects modelling to characterise the population averaged trajectory of the clinical signs and symptoms
- To carry out competing risk survival analysis to estimate the cumulative incidence of fever cleareance, hemataological recovery, and viral clearance (If data permitting)

Primary and Secondary Outcome Measures *(suggested maximum 200 words)*

The primary outcomes are the clinical signs and symptoms at baseline and during follow-up hospitalisation period in ETC

Proposed Methodology and Statistical Analysis Plan *(suggested maximum 400 words)*

Primary

Primary objective

In patients with a laboratory-confirmed EVD diagnosis, the frequency, relationships at baseline and temporal trends of clinical signs, symptoms, biomarkers, and the dynamics of viral load will be explored. Descriptive statistics (mean and standard deviation, geometric mean or median and range, and frequencies as appropriate), scatter plots, box plots and spghetti plots over time will be presented overall, within each country and at different time in the epidemic.

The evolution of the presence of different clinical signs and symptoms will be measured from the time of patient admission until death or discharge. The proportion of patients with given signs and symptoms will be presented at given time point during hospitalisation period (e.g: days 1, 2, 3 etc.).

The proportions will be presented by groups of interests, some of which (but not restricted to) are: age categories, gender, pregnancy, high viral load, country, different timepoints in the epidemics, duration of symptoms before admisison, intervention regimens.

Secondary objectives:

Longitudinal mixed effects modelling

Longitudinal mixed effects modelling with random slopes for patients will be considered to characterize the population averaged trajectory of the parameters of interest over time and variability between patients. Host factors will be used to explain this variability and group-based trajectories for geographic locations will be studied.

Competing risk survival analysis for symptomatic resolution

Data permitting, the time to symptomatic resolution of fever, anaemia and viral load (or any other symptoms) will be generated using competing risk survival analysis using cumulative incidence function (CIFs). In generation of CIFs, death or discharge will be considered as a competing risk endpoint. Differences in CIFs estimates between groups of interest will be tested using the Gray's *k*-sample test.

However, identification of putative factors associated with the symptomatic resolution will be carried out using Cox's regression analysis.

Ethics *(suggested maximum 300 words)*
Provide details of any ethical considerations relating to the research proposal.
Additionally, list any approvals required by your institution to undertake this work, list reference numbers of any approved proposals, or explain why no approvals are required.

This project involves accessing and re-using patient data collected in the context of a public health emergency for research. To ensure the ethical integrity of the project, the lead requestor's key responsibility is to protect the privacy and interests of the individuals and communities of data origin. This application will be using retrospective anonymised random data. This would ensure that the confidentiality of patients is maintained and avoid any outputs that could stigmatise individual populations.

The ethical clearance for this study has been obtained from the Guinean National Ethics Committee (CNERS), the Liberia National Research ethics Committee, the Sierra Leone National Ethics Committee and the Oxford Tropical Research Ethics Committee as a part of the Ebola Data Platform Research Agenda. This approval includes the project's detailed approach to data access, informed consent, vulnerable participants, protection of privacy, community engagement and benefits sharing, and will be renewed annually throughout the project.

No ethics committee review is required by AIMS to carry out research on anonymised data within the research mandate of the institution. Ebola s one of the diseases for which the insitution has a fundamental mandate to conduct research.

The project for which this application is made, adopts the data security measures put in place by the Ebola Data Platform. This is in compliance with general data protection regulation.

Publication and Dissemination Plan *(suggested maximum 300 words)*
Provide details of plans for authorship/acknowledgement of data contributors.
Provide details of timelines for publication and dissemination of research findings.

Results will be written up as essays as part of the requirements of the students' degree. It is anticipated that the findings will also be written up as a peer-reviewed publication (available open access), and posted on the IDDO website. We will ensure that results are made openly accessible and communicated to the Ebola research community.

The essence of this application is to build capacity, build awareness of how resources such as the Ebola Data Platform can strengthen researcher skills, promote access to data and

make available information on Ebola that would enable a robust response to any future outbreak.
To improve on access, many avenues will be explored to share the outputs of this work including but not limited to electronic, print and social media platforms.

Addressing Knowledge Gaps (*suggested maximum 300 words)*
Provide details of how this research will address knowledge gaps of importance to those affected by or at risk of emerging and poverty-related diseases.

The Ebola outbreak of 2013-2016 reinforced the need for collaboration and an integrated approach to the response to an outbreak of such magnitude. Since Ebola appeared in 1976 data and information has been scanty and not well coordinated. This project seeks to address that need by cross referencing different data sets and put forward a meaningful case definition for EVD that will influence policy and decision making.

There is currently a paucity of Information regarding best treatment options. The current analysis will help in characterisation of the signs and symptoms of patients with EVD as well as in identifying factors associated with outcomes.

This will go a long way in reducing case fatality rates and mortalities associated with EVD. Due to this paucity of information, the WHO Ebola/Marburg Research and Development Road map has called for the establishment of an interoperable system to enhance capabilities for collecting, reporting, analysing and sharing data. The EDP is a forum wherein several entities including the national health agencies and Ministries of health of Sierra Leone Guinea and Liberia and other institution like West Africa Health Organisation World health organization and Oxford University to name but a few are collaborating to pool resources in coming up with research questions of priority which the project for which this application is been made seeks to answer in an attempt to address the need identified by WHO.

Ebola-affected countries and its experts will guide the process in coming up with priority questions and the capacity needs that have to be addressed to ensure that a robust health system is created to respond to any future outbreaks. The results from this application would form the bedrock on which these recommendations to the various Government, institutions and Non- Government organizations will be made.
To the survivor cohort an explanation would be put forward for the improved case fatality and also triangulate the most prevalent complication as a sequela to their infection. Of concern to the survivors is their potential risk to their loved ones. Information on the prevalence of the virus or viral fragment of Ebola has now improved but still not yet optimal. Due to the large quantum of data from the EDP, an attempt will be made to help reform sexual and reproductive health guidelines for survivors.

The prognostic indicators as determined from this study will inform the management of future outbreaks of viral haemorrhagic fever in affected countries. It will place emphasis on which biomarkers that should be looked for to show progress of the disease and delineate specific interventions that would affect the health outcome. Hence it will trigger the evaluation of laboratory test and diagnostic required for Viral Haemorrhagic fever patients. This will then inform policy and guidelines on laboratory test done for Ebola patients.

Equity and Capacity Building *(suggested maximum 300 words)*
Provide details of how this research will support health equity and/or capacity building in endemic regions affected by or at risk of emerging and poverty-related diseases.
Please refer to the Ebola Data Platform Approaches to Capacity Building for guidance.

The project for which this application is made seeks to improve on the preparedness response and treatment of future outbreaks of Ebola Virus Disease. It is a collaborative effort between the National Public Health Institute of Liberia, the National Health Security Agency of Guinea, the African Institute of Mathematical Sciences (AIMS) and the University of Oxford. Importantly, the approach to evidence generation taken by this project maximises benefits to the communities affected by Ebola. Students at AIMS Cameroon will carry out statistical analysis and interpretation of findings, under the supervision of researchers from the University of Oxford and the National Public Health Institute of Liberia, which will contribute to capacity building in a region affected by infectious diseases. The skills acquired during the students' training will be used in the analysis of the data stored in repository of the EDP that may form the basis for future projects. As part of their training they will develop skills in analysing data of this type and in reporting their findings, which will help with future work on related research. Students will present and discuss their work with other members of AIMS. The project will also strengthen the collaboration between AIMS and the University of Oxford.

The strength of the Ebola case definition has also been questioned due to low sensitivity and specificity demonstrated in a number of settings. The project partners have prioritised the need to better understand how to balance the strength of the definition with the utility of the definition in different contexts for future outbreaks. By leveraging the pooling of the largest individual patient data repository assembled under the EDP, the most conclusive evidence on these research questions can now be generated by this project. This will be used to influence policy on future management of Ebola.

The equitable and sustainable partnership between the various partners is built on collaboration and data sharing to enhance capacity building and training opportunities for a more robust national health security system.

Funding *(suggested maximum 100 words)*
Provide details of how this research will be funded/resourced.

There is no specific funding for this project. The analysis done by the students is funded through AIMS. Participation of Trokon O. Yeabah is funded by a TDR Career Development Fellowship.

Scientific Review *(suggested maximum 200 words)*
Provide details of how the details of the project outlined above have been scientifically reviewed. This could be by your institution, a funder/donor or review committee.

The research priority as outlined in this application is part of the research questions in the Ebola data Platform research agenda. This application was peer-reviewed by Ben cooper a renowned statistician.

Data Variables

*Provide a list of the **data variables and data sources** required to achieve the research objectives.*
Note: Data sources can be listed as populations (e.g. all EVD-positive pregnant women, or all children under 16 years of age from Liberia) or as datasets from a source listed on the Accessing Data web page (these should be named by 'Contributing organisation, Country, City' as listed in the table). Get in touch if you have any questions about this ebolaDAC@iddo.org

Based on the case definition given by WHO which was applied in the outbreak of 2013 - 2016 which reads: 1) any person, alive or dead, who has (or had) sudden onset of high fever and contact with a suspected, probable, or confirmed EVD case-patient, or a dead or sick animal; or 2) any person with sudden onset of high fever and ≥3 signs/symptoms (headache, generalized or articular pain, intense fatigue, nausea/vomiting, loss of appetite, diarrhoea, abdominal pain, difficulty swallowing, difficulty breathing, hiccups, miscarriage); 3) unexplained bleeding; or 4) sudden unexplained death.

The data variables are:

Patient demographics

- Age
- Gender
- Weight
- Country of residence
- location of Ebola Treatment Centre
- Profession

Clinical signs and symptoms

- Abdominal pain
- Fever
- Diarrhea
- Vomitting
- Bleeding
- Confusion
- Unconsciouness
- Hypovolemic shock (dizzy)
- Anaemia
- Co-morbidity with malaria
- Body temperature
- Blood pressure

Disease diagnosis

- Laboratory test to confirm haemorrhagic fever diseases
- Laboratory test to confirm other diseases (HIV, malaria, tuberculosis)
- Date of sample collection
- Date of laboratory test

Disease exposure factors

- History of traveling to known hotspot area

- Direct contact with EVD suspected or confirmed case
- Participating in the funeral of an EVD suspected or confirmed case
- Touching the body of an EVD suspected or confirmed case
- Time between symptomatic onset to admission at ETC
- Time between admission at ETC and death/discharge

Disease specific characteristics

- Viral load (or cycle threshold value)
- Date of admission
- Date of onset of symptom
- Symptoms present at the time of admission (self-reported)
- Clinical signs (observed by a clinician)

Treatment (Debatable whether to include or not in a prognostic models)

- Antiviral/antibacterial/rehydratation (infusion or oral) therapy
- Duration and dosage of treatment administered
- Route of administration
- Date of treatment start
- Date of treatment completion

Laboratory parameters

- White blood cell count
- Haemoglobin
- Platelet
- C-reactive protein
- Electrolytes
- Creatinine
- Liver transaminases
- CPK
- Lactic acid
- Base excess
- PTT
- Antibody titers

Epidemiological demographics

- Time since Ebola outbreak per country
- The number of cases at that week per country