Ebola Data Platform - Data Access Application Form

Please review the Data Access Guidelines and the Data Transfer Agreement[1] 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	Primary Investigator
First name (given name)	Alicia
Surname (family name)	Genisca
Gender	Female
Position at employing organisation/ institution	Assistant Professor in Emergency Medicine & Pediatrics at the Warren Alpert Medical School of Brown University
ORCID ID https://orcid.org/	https://orcid.org/0000-0001-7679-1126
Email	[redacted]
Telephone/Skype/WhatsApp	[redacted]
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***1	
Institution name	Lifespan: Hasbro Children's Hospital/Rhode Island Hospital (Affiliate of Brown University Alpert Medical School)
Address	55 Claverick Street, 2nd Floor Providence, RI 02903
Department (if applicable)	

[1] 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.*

	Emergency Medicine	
Please acknowledge that your institution agrees to execute the Data Transfer Agreement (in the case of your application being approved)		YES

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	Ian Michelow, MD, MMed, DTM&H
1. Title	Associate Professor of Pediatrics
1. Organisation/Institution	Hasbro Children's Hospital/ Warren Alpert Medical School of Brown University
2. Name	Andres Colubri, MFA, PhD
2. Title	Computational Scientist
2. Organisation/Institution	Broad Institute of MIT & Harvard (until August 2020) University of Massachusetts Medical School (after August 2020) *of note, Dr. Colubri is also a co-applicant in the approved EDP project titled "Collaborative privacy: Derivation and validation of robust, and personalised clinical insights for Ebola Virus Disease."
3. Name	Tzu-Chun Chu, MPH
3. Title	Biostatician
3. Organisation/Institution	Brown University Center for Statistical Sciences
4. Name	Himanshu Vaishnav, MS
4. Title	Data Analyst
4. Organisation/Institution	Brown Emergency Medicine
5. Name	Adam Levine, MD, MPH
5. Title	Associate Professor of Emergency Medicine
5. Organisation/Institution	Rhode Island Hospital/ Warren Alpert Medical School of Brown University
SECTION B: RESEARCH PLAN	
Title of Proposed Research	Predicting Ebola Diagnosis and Mortality in Pediatric Patients
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.	No
Summary of Research in Lay Language *(suggested ~ 100 words)*	
Ebola Virus Disease (EVD) is a rare but severe illness. Although children represent a small number of the cases in many outbreaks, the disease case fatality is high, especially in children <5 years old. Creating a predictive model for diagnosis for pediatric patients can help to rapidly identify patients who need care. Additionally, a predictive prognostic model for pediatric mortality allows clinicians to allocate resources appropriately. There are few studies to date that address diagnosis and prognosis of EVD in pediatric patients.	

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

Ebola Virus disease (EVD) is a rare but severe illness that continues to plague many African countries. In 2014-2016, the largest EVD outbreak occurred affecting nearly 29,000 individuals and claiming over 11,000 lives[1]. The most recent outbreak is ongoing with over 3,400 individuals with EVD[2]. Early identification of pediatric patients who are EVD positive can lead to prompt initiation of treatment. Currently, pediatric patients are screened using the World Health Organization's (WHO) age-based criteria and held in Ebola Treatment Centers (ETC) while awaiting confirmatory testing. One challenge of these criteria is that they also apply to many common non-life-threatening febrile pediatric illnesses, placing children at risk for nosocomial transmission of EVD in ETCs and delaying management of their true illness. A diagnostic predictive model could help to risk stratify suspect cases of EVD, reducing exposure for EVD negative patients while expediting appropriate treatment for EVD positive patients. To date, there has only been one study that has endeavored to create a predictive model for diagnosis of EVD in pediatric patients[3]. This study provided useful information on risk factors for EVD in pediatric patients but has not been externally validated.

The case fatality ratio for EVD is high (40-90%). Although children represented a small portion of the total EVD positive cases in the 2014-2016 outbreak, they represent a much higher proportion of patients during the current outbreak with a high case fatality ratio, especially those < 5 years of age. No study has been conducted to create a reliable prognostic model for EVD mortality in pediatric patients. This is crucial for identification of pediatric patients who will need intensive monitoring and treatment as well as improving resource allocation within an ETC. Therefore, we propose conducting a large retrospective cohort study to develop new pediatric EVD diagnostic and prognostic models. These models can then be used to update current existing mHealth tools for rapid bedside diagnostic and prognostic evaluation of pediatric patients with EVD.

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

Objective 1: To derive and externally validate a new diagnostic prediction model for pediatric EVD. *Hypothesis 1:* The new diagnostic model will correctly identify EVD in pediatric patients potentially decreasing nosocomial EVD infection

Objective 2: To derive and externally validate a new prognostic prediction model for pediatric EVD. *Hypothesis 2:* A pediatric prognostic predictive model can identify factors that increase the risk of mortality from EVD in pediatric patients allowing health care providers to appropriately allocate resources for individual patients.

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

Primary outcome: The primary outcome of this study will be confirmed EVD for the diagnostic model and death from EVD for the prognostic model.

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

This study will use retrospective data collected on pediatric patients (<18 years old) for whom outcome data is available in the Ebola Data Platform. The derivation cohort will include patients presenting to the contributing ETCs during the 2014-2016 EVD outbreak in West Africa. The validation cohort will include patients presenting to the International Medical Corps (IMC) Mangina ETC in DRC in 2018-2019 and if available, data from other centers included in the Ebola Data Platform. We will initially study clinical and demographic variables from West African children likely to be associated with either the diagnosis of EVD (diagnostic model) or mortality

from EVD (prognostic model) based on our previous published experience[4] , a thorough review of the literature and input from experts in the field. An initial set of candidate variables will be analyzed using multivariable logistic regression to adjust for confounding. We will implement a forward stepwise regression algorithm to optimize model selection, which chooses the optimal number of main effects and interactions in the model. The final logistic regression models will then be converted into a scoring system using previously described methods[5].

Once derived, both the diagnostic and prognostic models will be externally validated using the data from DRC, including an assessment of both their discrimination and calibration. Model discrimination will be evaluated using the receiver operating characteristic curve (AUC) at consecutive threshold settings of the predicted probability. We will assess calibration by plotting the calibration curve across all predicted probabilities, which visualizes the model-based and observed probabilities. Other predictive indices including McFadden's pseudo-R2, Brier score, accuracy, sensitivity and specificity will also be presented to assess model performance.

We will assess whether the missing data are missing completely at random (MCAR) by comparing the characteristics between individuals with or without missing data. If not MCAR, we will use multiple imputation to address missing information, which reduces potential selection bias which may result from excluding these patients. Multiple imputation will be carried out using the aregImpute function from the R package Hmisc, which uses bootstrap resampling to resemble a Bayesian predictive distribution from observed data, and obtain N imputed datasets. The flexible additive model will be fitted on different bootstrapped resamples to predict values for observations with missing values in the ith imputation. We will begin by setting N=20, and increase the number of imputations as needed.

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 research involves only a retrospective chart review of routinely collected data. As such, we do not plan to obtain informed consent from patients. We requested exemption of informed consent for this purpose, as this research poses no risk to participants, and because all data are de-identified and routinely collected. Approval of exemption has been provided by the Lifespan Institutional Review Board (# 1527875-2, approved 1/27/2020).

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.

Authorship and acknowledgement of data contributors will follow the International Committee of Medical Journal Editors standards.

Our goal is to have a publishable manuscript of the predictive models and scoring systems within 6 months of receipt of the data with eventual publication 6 months to a year afterwards.

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.
In keeping with the WHO's Research and Development Roadmap for Ebola, the IDDO has set an agenda of research priorities for EVD including understanding host factors that contribute to EVD diagnosis and mortality, risk stratification and optimal supportive therapies. With the data from the Ebola Data Platform, our research proposes to be the largest study of the pediatric population affected by EVD. Development of a diagnostic prediction model for pediatric EVD can assist in appropriate triage of potential pediatric EVD cases, reducing exposure of EVD negative children to EVD and more rapidly getting EVD positive children into optimal care. Additionally, the development of a prognostic prediction model for pediatric EVD mortality can assist clinicians in allocating resources appropriately, especially in resource-limited settings where Ebola epidemics are most likely to occur. With the recent development of new therapies for EVD, tools that can rapidly identify children with EVD who might benefit from these treatments are more important than ever.
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.*
Per IDDO guidelines, our research will support health equity by fulfilling the first objective of identifying and responding to the EVD research priorities of countries affected by or at risk of EVD outbreaks. Additionally, in terms of capacity building, our models can be used to update existing mHealth tools to enhance real time diagnostic and prognostic evaluations performed by bedside clinicians.
Funding *(suggested maximum 100 words)* *Provide details of how this research will be funded/resourced.*
This research is funded by the Rhode Island Foundation (Grant #: 5222_20200596). Award amount $25,000. 4/2020 – 6/2021.
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.*
This research study has reviewed by the Rhode Island Foundation whose goal is to improve healthcare. This study has been scientifically reviewed by a diverse group clinicians and statisticians in addition to being approved by the Lifespan Institutional Review Board.
SECTION C: DATA
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*
All available variables for pediatric patients (<18 years of age) presenting to contributing ETUs in the Ebola Data Platform. If we are unable to have access to all variables, we would like to request the following:

- Demographic variables: age, sex, height, weight, ETU location, Ebola contact, immunization status, caregiver present
- Vital signs: temperature, heart rate, respiratory rate, oxygen saturation, blood pressure
- Symptoms at triage and during ETU stay (fever, headache, emesis, bloody vomit, diarrhea, bloody diarrhea, anorexia, cough, joint pain, muscle pain, bone pain, chest pain, abdominal pain, hiccups, gum hemorrhage, nasal hemorrhage, other hemorrhage, difficulty swallowing, confusion, coma, conjunctivitis, photophobia, dyspnea, dysphagia, weakness, rash
- EVD outcomes: positive or negative for EVD, final diagnosis if negative for EVD, PCR cycle times, length of stay in ETC, survival or death, disposition from ETC
- Other diagnoses: parasitic, bacterial or viral infection
- Laboratory studies, if available: complete blood count, metabolic panel, C-reactive protein, erythrocyte sedimentation rate, cultures
- Medical treatments: intravenous fluids (type and amount), oral fluids, antimalarials, antibiotics, antivirals, vitamin supplementation, antipyretics, antispasmodic, anti-nausea, vaccine administration

References:

1. WHO. Ebola Situation Report - 30 March 2016. Available from: http://apps.who.int/ebola/current-situation/ebola-situation-report-30-march-2016.
2. WHO. Ebola Virus Disease Democratic Republic of Congo: External Situation Report – 24 June 2020. Available from: https://www.who.int/publications/i/item/10665-332654.
3. Fitzgerald F, Wing K, Naveed A, et al. Development of a Pediatric Ebola Predictive Score, Sierra Leone. Emerg Infect Dis. 2018; 24(2): 311-319.
4. Smit MA, Michelow IC, Glavis-Bloom J, et al. Characteristics and Outcomes of Pediatric Patients With Ebola Virus Disease Admitted to Treatment Units in Liberia and Sierra Leone: A Retrospective Cohort Study. Clin Infect Dis. 2017; 64(3):243–9.
5. Levine AC, Shetty PP, Burbach R, et al. Derivation and Internal Validation of the Ebola Prediction Score for Risk Stratification of Patients With Suspected Ebola Virus Disease. Ann Emerg Med. 2015; 66(3): 205-293.