Title: Prognostic models for Ebola virus disease derived from data collected at five treatment units in Sierra Leone and Liberia: performance, external validation, and risk visualization

Authors: Andres Colubri[1,4,5*], Adam C. Levine[2,3*], Mathew Siakor[3], Vanessa Wolfman[3], Mary-Anne Hartley[7,8], Tom Sesay[9], August Felix[4], Pardis C. Sabeti[1,4,5,6]

Affiliations:

[1] Harvard University, Department of Organismic and Evolutionary Biology, Cambridge, MA, USA

[2] Brown University, Warren Alpert School of Medicine. Providence, RI, USA

[3] International Medical Corps. Los Angeles, CA, USA

[4] Broad Institute of MIT and Harvard, Cambridge, MA, USA

[5] Howard Hughes Medical Institute, Chevy Chase, MD, USA

[6] Harvard School of Public Health, Boston, MA, USA

[7] University of Lausanne, Faculty of Biology and Medicine, Switzerland

[8] GOAL Global, Dublin, Ireland

[9] Sierra Leone Ministry of Health and Sanitation

Abstract

Background: The 2014-2016 Ebola Virus Disease (EVD) outbreak highlighted the need for rigorous, rapid, and field-deployable tools to enable case management. We previously introduced an approach for EVD prognosis prediction, using models that can be implemented in the field and updated in light of new data. Here we enhance and validate our methods with the largest published EVD dataset to date. We also present a proof-of-concept medical app that summarizes patient information and offers tailored treatment options using an interactive risk visualization for quick interpretation and decision-making.

Methods and Findings: We derived prognosis prediction models for EVD using data from 470 patients admitted to five Ebola treatment units (ETUs) operated by International Medical Corps (IMC) in Liberia and Sierra Leone. We fitted logistic regression models, handled missing data by multiple imputation, and conducted internal validation with bootstrap sampling. We also validated our models with independent datasets from two treatment centers in Sierra Leone comprising 106 patients at Kenema Government Hospital and 158 patients at the GOAL-Mathaska ETU in Port Loko district. We corroborated earlier reports on the importance of viral load and age as mortality predictors and identified jaundice and bleeding to be features with highest predictive value at presentation. Additional clinical symptoms at presentation, although individually weakly correlated with outcome, help broaden sensitivity and refine discrimination of the models. The app provides a visual representation of the predictive outcome as well as attributing clinical protocols adjusted by demographic parameters and prioritized to target the largest contributing factor to overall risk. The app is freely available under the name of Ebola RISK on Google Play and Apple's App Store.

Conclusions: We derived and validated high performance models of EVD prognosis prediction from the largest and most geographically diverse EVD patients available to date. The performance was maintained during external validations on two independent datasets representing different treatment settings and mortality rates, which suggests that the models could be generalized to new populations. These models and derived tools may better inform treatment choices in future EVD outbreaks. The risk visualization app also provides a template to validate additional datasets used in developing novel clinical-decision support systems for EVD and other emerging infectious diseases.

Keywords: Ebola Virus Disease, Prognosis, Logistic Model, Risk Visualization, Medical Apps

Introduction

The 2014-2016 outbreak of EVD caused a worldwide health crisis with more than 28,000 cases and 11,000 deaths, the vast majority of which occurred in the West African countries of Liberia, Sierra Leone, and Guinea. Despite its notoriety as a deadly disease, the pathology EVD includes a range of outcomes, spanning from asymptomatic infection to complex organ failure, with mortality rates of under 20% achievable in high-income countries where extensive resources can be applied on the few cases that were treated there. Clinical care of highly contagious diseases such as EVD in remote and low-resource settings is far more challenging, hindered by limited availability of trained personnel, restricted time that can be allocated to each patient due to difficult-to-wear personal protective equipment, and lack of supplies. Prioritizing time and material resources for high-risk patients is one approach to decrease overall mortality when subject to such constraints (1). A complementary approach is to use tools providing clinical instructions for management, training, and improved protocol adherence (2, 3).

We previously introduced the use of prognostic models that can be deployed on medical apps for the purpose of risk stratification in EVD (4). Prognostic models can enable the early identification and triage of high-risk patients could be useful in low-resource areas to better allocate supportive care. For example, physicians could more frequently monitor those patients at increased risk and decide between standard and more aggressive therapy. Our original models were developed on just the single publicly available dataset at that time by Schieffelin et al. (5), which includes 106 Ebola-positive patients at Kenema Government Hospital (KGH). These models, which we validated internally, gave good results by outperforming simpler risk scores and allowing to choose

from different set of predictors depending on the available clinical data, showing the potential for this approach. However, given this was a single study site in one country, with a small patient cohort from one period during the outbreak and affected by a large incidence of missing data, questions remain about the generalizability of the approach.

We sought to create new prognostic models by using the IMC EVD patient cohort, the largest and most diverse available to date (5-10). It is comprised of 470 confirmed cases from five Ebola Treatment Units (ETUs) in Sierra Leone and Liberia, admitted between September 2014 and September 2015. Given the larger sample size and diversity in patient origin, we can expect to generate models that are not overfitted to the characteristics of a specific patient group, and that can be generalized to new EVD cases. The IMC dataset includes demographic information and clinical signs/symptoms of patients at presentation, RT-PCR Cycle Threshold (CT) measurements (quantifying viral load) done at admission and approaching discharge, daily updates on their signs/symptoms, and overall wellness assessments. The clinical and lab protocols were consistent across the five ETUs, making it possible to aggregate individuals into a single cohort.

External validation across sites is critical to establish the geographic and demographic range to which the models may generalized (11). Ideally, the model should be applied to a dataset that was obtained independently from the cases originally used for model training, but even then, porting prognostic models from one center to another is challenging (12). To this end, we report two independent external validations on datasets collected at different health care centers with independent patient catchment areas on patients reporting at different time points of the epidemic. The first, includes the 106 Ebola-positive patients at KGH described by Schieffelin et al. and

collected in the first months of the outbreak. The second, described by Hartley et al. (13), comprised 158 Ebola patients who were treated in an ETU run by GOAL global during the final months of the epidemic under conditions that should better represent future outbreak responses.

The ultimate goal of these models is to aid clinical management decisions on ground, for which we introduced a mobile app for deployment. These apps can offer unprecedented convenience and precision at the point of care (14, 15). Platforms for clinical data collection, such as CommCareHQ (https://www.commcarehq.org), REDCap (https://www.project-redcap.org/), and Open Data Kit (https://opendatakit.org), are in use on low-cost smartphones by front-line workers in several low-income countries around the world (16-21). Motivated by these platforms, we implemented our models into a freely available medical app for Android devices. We choose the Android operating system (OS) due to the growing adoption of smartphones using that OS across Africa (22). Internet connectivity is still limited, particularly in rural areas, and so we decided against implementing the models as a web-based app, even though that route would make it more widely available. Once installed on a smartphone or tablet, our app does not require any further internet access, and it includes treatment information from WHO guidelines for Ebola and other viral hemorrhagic fevers (23, 24), structured around individualized risk predictions for faster and better-targeted access.

Methods

IMC Patient Cohort

The cohort used to develop the prognostic models in this study includes patient data collected at five ETUs operated by IMC in Liberia and Sierra Leone between September 15, 2014 and September 15, 2015. The ETUs were located at Lunsar (Port Loko District), Kambia (Kambia District), and Makeni (Bombali District) in Sierra Leone, and at Suakoko (Bong County) and Kakata (Margibi County) in Liberia. The majority of the patients did not come from holding units and presented directly to the IMC ETUs, with an overall Case Fatality Rate (CFR) across the 5 ETUs of 58%. Collection and archival protocols are detailed in Roshania et al (25). The Sierra Leone Ethics and Scientific Review Committee, the University of Liberia – Pacific Institute for Research & Evaluation Institutional Review Board, the Lifespan (Rhode Island Hospital) Institutional Review Board, and the Harvard Committee on the Use of Human Subjects provided ethical approval for this study and exemption from informed consent. A data sharing agreement was approved by IMC and the Broad Institute, following IMC's Research Review Committee Guidelines (available online at https://internationalmedicalcorps.org/document.doc?id=800).

Data collection

Trained nurses, physician assistants, physicians, and psychosocial support staff recorded patient demographic, clinical, and support data at least daily from admission to discharge on standardized paper forms – as part of routine clinical care and for epidemiologic purposes. Local data officers entered this data into separate electronic databases at each ETU, which were combined together into a unified database. The RT-PCR data were obtained from four laboratories. The United States

Naval Medical Research Center Mobile Laboratory in Bong County, Liberia, served the Bong and Margibi ETUs; the Public Health England (PHE) labs in Port Loko and Makeni in Sierra Leone processed samples from Lunsar and Makeni; the Nigerian Lab served the Kambia ETU.

Exploratory and Univariate analysis

The primary variable of interest for patients admitted to the ETUs was final disposition (survived, deceased, or transferred). We constructed the Cycle Threshold (CT) variable using the values from the PCR drawn on admission, or from the second PCR draw when the first was missing (in 155 cases), and the second draw took place no more than two days after admission. The CT value is an inversely proportional proxy of viral load, with a cut-off of 40 cycles considered as negative. We used the visual exploration tool *Mirador* (https://fathom.info/mirador/) to examine correlation between disposition and all the explanatory variables available when the patient was admitted to the ETU, including demographic, triage, rounding, outcome, and lab data (CT.) We carried out an initial univariate analysis of all factors against disposition, using the χ^2 test with Yates correction for the binary variables, and the point biserial correlation test for numerical variables.

Logistic Regression with Multiple Imputation

We generated several logistic regression models using predictors available at presentation, following the pre-specified protocol described here, which is based on the model-building steps recommended by Harrell (26). In order to limit overfitting, we applied the commonly accepted heuristic of keeping the number of Degrees of Freedom (DOFs) in our models below M_{max}=N/15 (27), where N is the minimum count over the two disposition categories (survived or deceased;

Transferred was recoded as missing, but this affected only 4 cases.) In the entire dataset, N=197, so M_{max}~13.

In order to limit the number of predictor variables to 13, we removed variables with high incidence of missing values (such as confusion and coma); and conducted redundancy analysis by predicting each variable as a function of the rest (excepting disposition) with the function redun() from the Hmisc package available in R, and removed variables that could be predicted with an R^2 higher than 0.2. Previous studies (5-10, 28) also informed variable selection. We grouped variables by performing hierarchical clustering with R's varclus() function, using the pairwise Hoeffding D statistic as similarity measure, selecting one variable from each cluster for inclusion into the final models. Patient age and body temperature exhibit non-linear dependencies with disposition. We modeled them with restricted cubic splines (RCS) with three knots each at 5, 10, and 30 years for age, and 35, 37, and 40 °C for temperature (Suppl. Fig S1). RCS were not considered for the CT term since visual inspection of the CFR as function of CT suggested that a linear term would be sufficient to represent this dependency (Suppl. Fig S2).

We performed standard logistic regression in R, using the rcs() function from the Hmisc package to handle the RCS terms. We applied multiple imputation with the MICE package to complete missing records and tested the Missing Completely At Random (MCAR) condition with the MissMech package. Temperature was missing in 56% of the cases, so we added the binary variable Fever, missing only one value, to the imputation procedure with the aim of producing better imputations for temperature. We generated 50 imputed datasets and fitted each one separately. We pooled the resulting 50 fitted models into a single average model, which we used to make

predictions on new data. Although this inverts the formally correct order of predicting with each imputed model first and then averaging the results, simulation studies show that both approaches yield comparable results (29). We conducted bootstrapping validation (30) by repeatedly training each model on 1000 bootstrap replicates of each of the 50 imputed datasets, and calculating various optimism-corrected metrics (31): Area Under the Curve (AUC), overall accuracy, Brier score, calibration error (32), and adjusted McFadden pseudo-R^2. We applied Fisher's transformation (33) to estimate the means and confidence intervals (CIs) of the statistics over the multiple imputations. We generated the Receiver Operating Characteristic (ROC) and calibration curves by merging all the predictions from each bootstrap, and then averaging over the imputations.

Finally, we used the regression models to stratify the patients into risk groups. We defined low, medium, and high-risk groups based on the predicted risk (<0.3 for low, 0.3-6 for medium, and >0.6 for high) so that each group is sufficiently populated to be clinically meaningful (low-risk=6.6% of patient cohort, medium-risk= 20%, and high-risk=75%). All these calculations were carried out using only complete data, since in the training set we still have a significant percentage of patients with complete records (more than 100), and we sought to minimize the effect of multiple imputation on the predictive performance of the models due to higher error and variance.

External validation

We did two external validations on independently collected datasets from Sierra Leone. The KGH dataset described by Schieffelin et al. (5) is the only such database to be made publically available at the time of this study (https://dataverse.harvard.edu/dataverse/ebola). It includes 106 EVD-

positive cases treated at KGH between May 25 and June 18, 2014. CFR among these patients was 73%. Sign and symptom data were obtained at time of presentation on 44 patients who were admitted and had a clinical chart. Viral load was determined in 58 cases. Both sign and symptom data and viral load were available for 32 cases. We generated 50 multiple imputations with MICE to apply the IMC models on the KGH cases with incomplete data. The GOAL dataset described by Hartley et al. (13, 34) includes 158 EVD-positive cases treated at the GOAL-Mathaska ETU in Port Loko between December 2014 and June 2015, where the CFR was 60%. Ebola-specific RT-PCR results and detailed sign and symptom data was available for all 158 patients. The Ebola-specific RT-PCRs recorded in the GOAL dataset were performed by the same PHE laboratory system as for the majority of the Sierra Leonean IMC data. Average CT values reported in this dataset between survival and fatal outcomes were not statistically different from that recorded by the IMC.

The KGH dataset includes RT-PCR data as viral load (VL) quantities expressed in copies/ml, but the corresponding CT values are no longer available. Since the IMC models use CT as a predictor, we transformed log(VL) to CT by solving for the standard qPCR curve transformation $\log(VL) = m \times CT + c0$, such that the minimum VL in the KGH dataset corresponds to the maximum CT in the IMC dataset, and vice versa. The assays used for diagnosing patients at KGH and IMC have very similar limits of detection (35-37), which justifies our VL-to-CT transformation. We also note that a ≈10-fold increase in Ebola VL corresponds to a 3-point decrease in CT (38). Based on this relationship, $-3/m$ in our formula should be close to 1, which is indeed the case ($-3/m=0.976$ using the m and c0 constants derived from the KGH and IMC data).

Risk visualization on mobile apps

We developed a mobile app for Android mobile devices that integrates patient data with the prognostic models and a custom risk visualization. This app only requires internet connectivity to be installed the first time, and it can be used even when the device is offline afterwards. This is an important consideration as health care workers, the intended users of this app, are often deployed in rural or remote locations with limited internet access. The choice of the Android OS was also informed by the increasing adoption of affordable Android smartphones in low and medium income countries in Africa and elsewhere. Users can enter basic information (age, weight), clinical signs and symptoms, and lab data of a patient obtained at triage into the app; the app then computes the risk score by selecting the appropriate model for the available indicators and offers two different visualizations of the death risk of the patient. In the first visualization, the numerical value of the risk score is shown at the top, and the magnitude of the contributions to the final score from each term in the model are represented graphically below using the patient-specific charts from by Van Belle and Van Calster, designed to visualize logistic regression predictions (39). In these charts, each contribution determines the length of a bar alongside a horizontal line whose total extension represents the maximum contribution observed in the training data.

The second risk visualization in the app displays the contributions to the patient risk with less detail but provides an entry point to further information of clinical relevance. The risk score is presented using a discrete scale: low, moderate, and high, and the predictors that contribute to the score above a configurable threshold are arranged in a list, ranked by decreasing contribution. When the user selects any of the predictors from the list, corresponding to a specific sign/symptom or laboratory result, the app shows treatment guidelines addressing the condition associated to the

selected predictor. There guidelines were manually curated from the WHO manuals for clinical management of patients with Ebola virus disease (23) or other viral hemorrhagic fever (24) in care units/community care centers and then incorporated into the app. Tables describing dosage of various medications and drugs (e.g. rehydration solution and antimalarials) as functions of age/weight have the appropriate entry highlighted according to the patient's information.

Results

1. Prognostic potential and prevalence of signs and symptoms recorded at triage

Triage symptoms reported by over 50% of fatal Ebola patients were anorexia/loss of appetite, fever, weakness, musculoskeletal pain, headache and diarrhea (Fig 1A and Table 1A). The prevalence of several triage symptoms was notably different between fatal and non-fatal outcomes, as can be seen by comparing their ranking (Fig 1A) or their differential prevalence (Fig 1B). However, few variables were significantly associated with patient outcome, suggesting that most clinical signs and symptoms have little predictive ability on their own, at least when considered at triage alone. Only CT, age (Table 1B), and jaundice (Table 1A) were associated at a level of $P<0.05$, while red eyes, confusion, breathlessness, headache, and bleeding were weakly associated at $P<0.15$. However, statistical association of the variables when taken alone might be due to confounding effects in the data. We thus set out to investigate the performance of individual variables within the context of specific multivariate models.

Table 1: Univariate analysis

Table 1A

Variable	*Total %*	*Non-fatal %*	*Fatal %*	*Missing %*	*OR 95% CI*	*P-value*
Jaundice	24/464 (5)	4/197 (2)	20/267 (7)	1/470 (0)	3.91 (1.31, 11.62)	0.016
Red eyes	128/464 (27)	64/197 (32)	64/267 (23)	1/470 (0)	0.66(0.43, 0.99)	0.054
Coma	5/178 (2)	0/83 (0)	5/95 (5)	292/470 (62)	NA	0.096
Confusion	16/178 (8)	4/83 (4)	12/95 (12)	292/470 (62)	2.86 (0.88, 9.23)	0.120
Breathlessness	109/464 (23)	39/197 (19)	70/267 (26)	1/470 (0)	1.44 (0.92, 2.24)	0.133
Headache	268/464 (57)	122/197 (61)	146/267 (54)	1/470 (0)	0.74 (0.51, 1.08)	0.142
Bleeding	26/464 (5)	7/197 (3)	19/267 (7)	1/470 (0)	2.08 (0.86, 5.05)	0.148
Asthenia/Weakness	334/464 (71)	135/197 (68)	199/267 (74)	1/470 (0)	1.34 (0.89, 2.02)	0.187
Diarrhea	234/430 (54)	96/187 (51)	138/243 (56)	35/470 (7)	1.25 (0.85, 1.83)	0.304
Malaria	49/225 (21)	17/94 (18)	32/131 (24)	241/470 (51)	1.46 (0.76, 2.83)	0.331

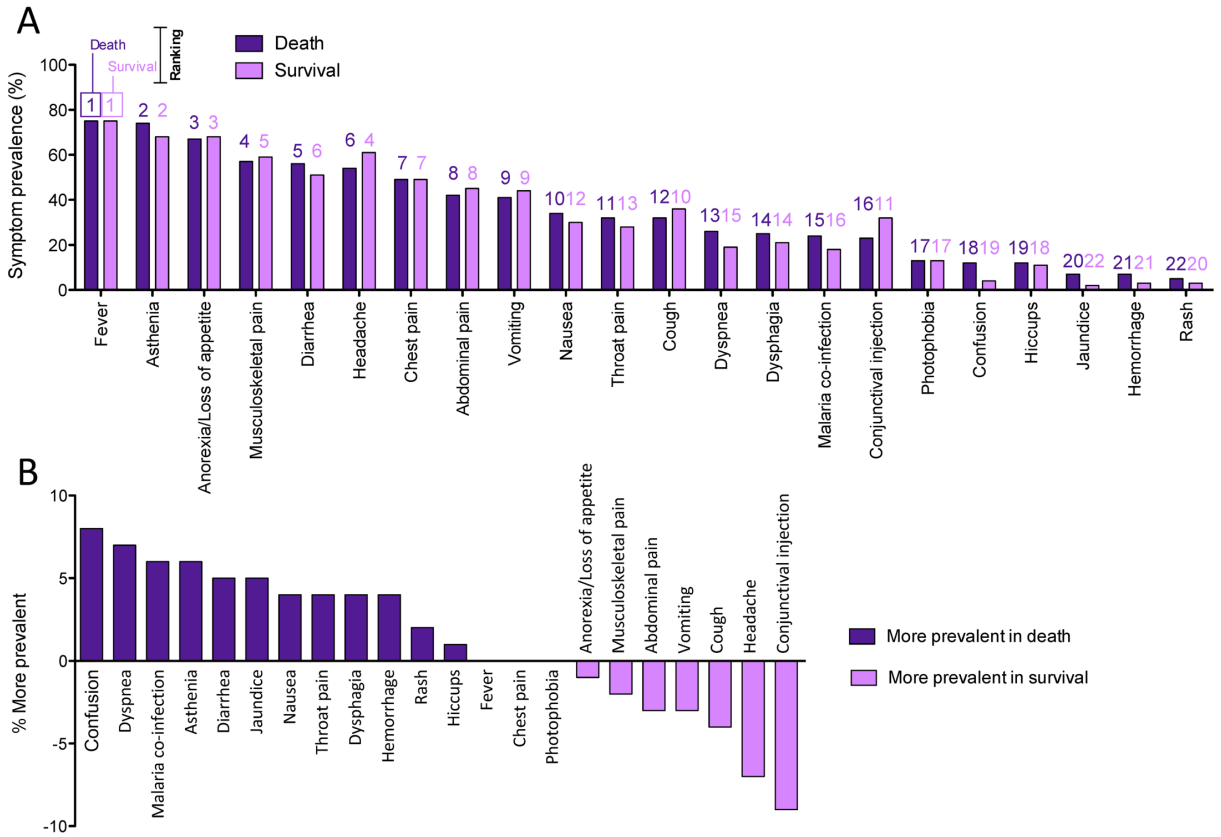
Swallowing Problems	112/464 (24)	43/197 (21)	69/267 (25)	1/470 (0)	1.25 (0.81, 1.93)	0.374
Vomiting	197/464 (42)	87/197 (44)	110/267 (41)	1/470 (0)	0.89 (0.61, 1.29)	0.587
Nausea	94/286 (32)	35/114 (30)	59/172 (34)	179/470 (38)	1.18 (0.71, 1.96)	0.613
Abdominal Pain	203/464 (43)	89/197 (45)	114/267 (42)	1/470 (0)	0.90 (0.62, 1.31)	0.662
Bone/Muscle/Joint Pain	272/465 (58)	118/197 (59)	154/268 (57)	0/470 (0)	0.90 (0.62, 1.31)	0.666
Throat Pain	55/178 (30)	24/83 (28)	31/95 (32)	292/470 (62)	1.19 (0.63, 2.26)	0.709
Cough	61/178 (34)	30/83 (36)	31/95 (32)	292/470 (62)	0.86 (0.46, 1.59)	0.738
Hiccups	55/464 (11)	22/197 (11)	33/267 (12)	1/470 (0)	1.12 (0.63, 1.99)	0.805
Rash	8/178 (4)	3/83 (3)	5/95 (5)	292/470 (62)	1.48 (0.34, 6.40)	0.867
Chest Pain	88/178 (49)	41/83 (49)	47/95 (49)	292/470 (62)	1.00 (0.56, 1.81)	0.889
Photophobia	24/178 (13)	11/83 (13)	13/95 (13)	292/470 (62)	1.04 (0.44, 2.46)	0.892
Anorexia/ Loss of Appetite	316/465 (67)	135/197 (68)	181/268 (67)	0/470 (0)	0.95 (0.64, 1.42)	0.900
Fever	349/464 (75)	148/197 (75)	201/267 (75)	1/470 (0)	1.01 (0.66, 1.54)	0.944

Table 1B

Variable	*Mean non-fatal cases (95% CI)*	*Mean fatal cases (95% CI)*	*Missing fraction (%)*	*Pearson's R*	*Odds-Ratio (95% CI)*	*P-value*
Cycle Threshold	26.72 (15.92, 37.52)	22.23 (11.18, 33.28)	29	-0.371	0.331 (0.23, 0.47)	<0.0001
Patient Age	28.49 (0.00, 58.72)	32.03 (0.00, 72.10)	0	0.095	1.326 (1.01, 1.74)	0.043
Body Temperature	37.41 (35.50, 39.32)	37.67 (35.34, 40.01)	56	0.116	1.391 (0.94, 2.06)	0.099
Days of Fever	3.44 (0.00, 7.74)	3.56 (0.00, 7.90)	74	0.025	1.048 (0.74, 1.48)	0.79

Correlation between either binary (A) or continuous (B) clinical variables and the outcome of death. Marginal odds-ratios were obtained from the univariate logistic regression model for death using each variable alone as a predictor. For continuous variables, the Pearson's R correlation coefficient is used and the odd-ratios correspond to inter-quartile range changes in the predictor

Figure 1: Prevalence of clinical signs and symptoms recorded at triage.



(A) Prevalence of clinical characteristics at triage amongst Ebola patients who either survived or died, ranked according to the prevalence in fatal outcomes. Rankings from 1–22 are listed above each bar: purple for the outcome of death and pink for survival. (B) Differences in symptom prevalence between EVD survivors and those who died. Positive values are more prevalent in fatal outcomes. Negative values are more prevalent in survivors.

2. Derivation and performance of multivariate models

Our key goal was to derive prognostic models from the IMC dataset with as much detail as possible, while limiting overfit to the training data. For this purpose, we constructed a "full" model comprising a maximal (DOF=13) but non-redundant set of variables. This model, obtained with

our variable selection strategy, included age of patient, initial CT and temperature, jaundice, bleeding, asthenia/weakness, vomiting, diarrhea, headache, and abdominal pain at presentation (Table 2A). Jaundice and bleeding were significant (P=0.021 and P=0.046, respectively), while the other sign/symptoms were still non-significant at the 0.05 level.

In order to define a baseline performance level, we also fitted a "minimal" model (Table 2B) including only CT and age, since these are the strongest predictors of outcome on their own, as observed in our data and reported by other researchers (10). Examination of the coefficients' P-values shows that CT and age are highly significant at P<0.0001 in both models.

Table 2: Full and Minimal model descriptions

Table 2A

Variable	*Coefficient (95% CI)*	*Odds-Ratio*	*P-value*
Cycle Threshold	-0.152 (-0.204, -0.101)	0.38	<0.0001
Patient Age	-0.044 (-0.078, -0.010)	0.34	0.011
Patient Age'	0.088 (0.044, 0.132)	3.46	0.0001
Body Temperature	-0.126 (-0.883, 0.631)	0.82	0.74
Body Temperature'	0.628 (-0.600, 1.856)	1.99	0.31
Headache	-0.128 (-0.606, 0.351)	0.88	0.6
Bleeding	1.112 (0.035, 2.189)	3.04	0.043
Diarrhea	0.082 (-0.435, 0.599)	1.09	0.76
Jaundice	1.516 (0.230, 2.802)	4.56	0.021
Vomit	-0.197 (-0.701, 0.308)	0.82	0.44
Abdominal Pain	-0.053 (-0.547, 0.441)	0.95	0.83
Asthenia/Weakness	0.167 (-0.362, 0.696)	1.18	0.53

Table 2B

Variable	*Coefficient (95% CI)*	*Odds-Ratio*	*P-value*
Cycle Threshold	-0.163 (-0.213, -0.114)	0.28	<0.0001
Patient Age	-0.051 (-0.083, -0.019)	0.25	0.0019
Patient Age'	0.096 (0.053, 0.138)	4.41	<0.0001

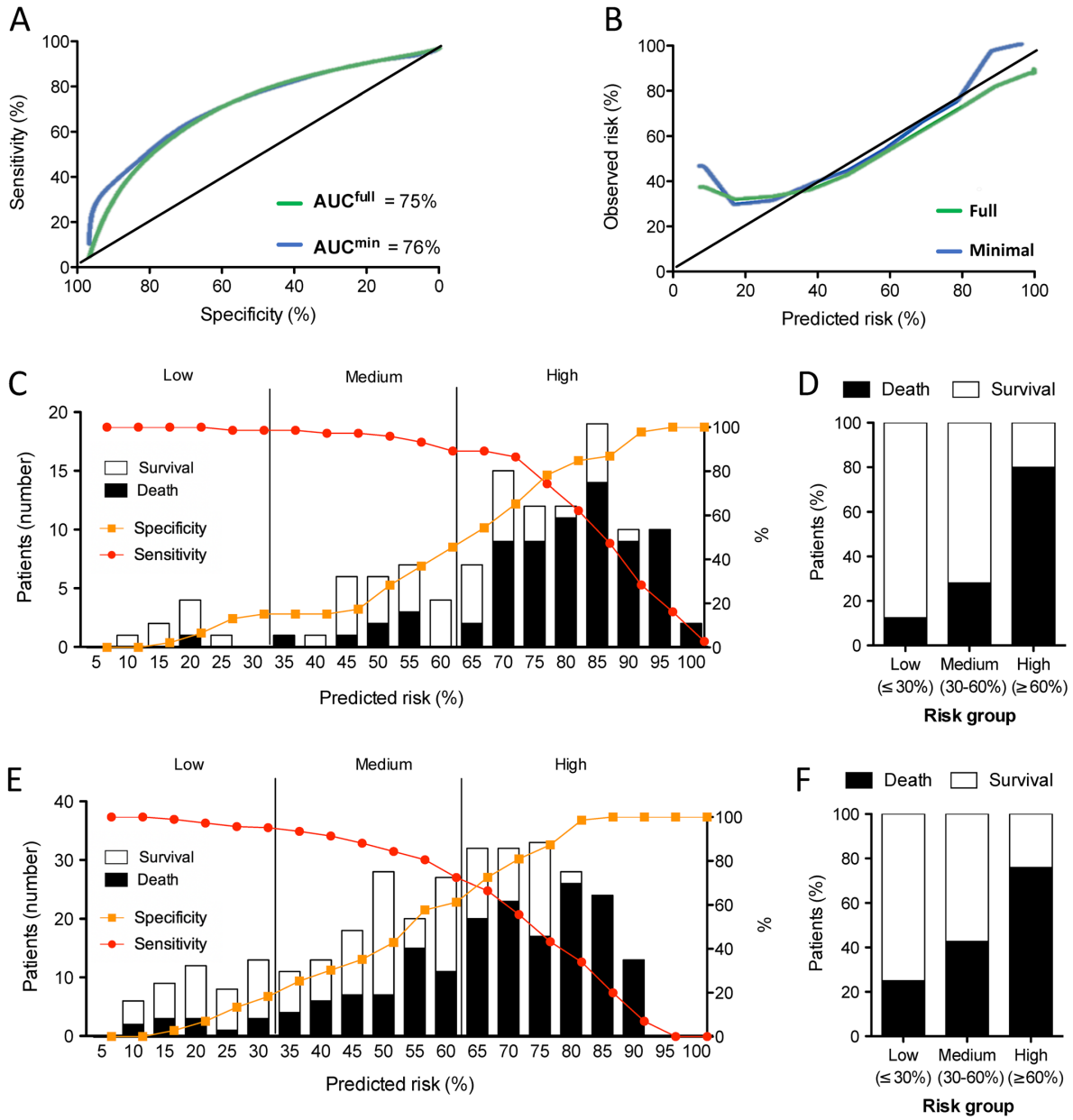
Multivariate logistic regression for full (A) and minimal (B) models. Coefficients are shown with their 95% confidence intervals as well as the corresponding odds-ratios, and P-values. The odds-ratios in continuous variables indicate the change in mortality by one interquartile range increase in the value of the variable. Variables with an apostrophe at the end of the name indicate the non-linear contribution to the corresponding Restricted Cubic Spline (RCS) term.

Predictive performance is similar between the full and minimal models (Table 3A), when initially evaluated on the training set. This evaluation includes optimism-correction, as described in the methods, to account for the fact that performance is likely to be overestimated on the training set. Discrimination, or the ability to distinguish between different outcomes, was quantified with the AUC statistic. The 95% CIs for the AUC of the minimal and full models are essentially overlapping at (0.7, 0.8), and the corresponding ROC curves are virtually indistinguishable (Fig 2A). Both models exhibit similar calibration, a measure of agreement between the predicted and observed mortality risks (Fig 2B). The calibration error is slightly lower in the full model, 0.018 vs 0.019, but the 95% CIs are also overlapping. However, the adjusted R^2 score indicates that the full model, corrected by model size, is a better fit for the data with a CI of (0.179, 0.276), against (0.133, 0.217) for the minimal model.

Even though overall performance as measured by AUC and calibration is similar between the minimal and full models, we found that the full model could lead to better patient stratification. The full model results in larger differences in observed mortality between the tree groups (Fig 2D vs 2F), with CFR nearing 10% in the low risk group, 30% in the medium risk group, and over 80% in the high-risk group. In contrast, when the risk groups are defined using the minimal model, CFR in the low risk group is slightly above 20%, 40% in the medium risk group, and below 80% for the high-risk group.

By the measures presented thus far, the full and minimal models are largely equivalent, at least on the training data. The edge by the full model in terms of stratification could be expected since it includes a larger set of predictors, but a more definitive validation of the models on the two independently-obtained test datasets is presented in the next section.

Figure 2: Performance of prognostic models in the IMC training dataset



ROC (A) and calibration (B) curves for the two prognostic models (full and minimal) using the bootstrap samples taken from the training data. The sensitivity and specificity of predicting mortality in Ebola patients using the full (C) and minimal (E) models in the IMC training dataset. Sensitivity (red) and specificity (orange) are plotted according to the risk prediction of each model.

Prevalence of survivors and those with fatal outcome are displayed as bar graphs and risk category cut-offs are shown as vertical lines. Percentage of survivors and patients with fatal outcome classified in each risk category for the full (D) and minimal (F) models. Graphs C and D represent only complete records according to the parameters of the full model (120/470, 26%). Graphs E and F represent only complete records according to the parameters of the minimal model (327/470, 70%).

3. External validation

External validation on the KGH dataset shows (Table 3B) that the full model is able to accurately predict outcome for 75% of the patients, versus 67% in the minimal model. The ROC curve of the full model is nearly perfect (Fig 3A), even though the number of KGH cases with complete PCR and signs/symptoms data is only 32. External validation on the imputed KGH data, consisting of 106 cases, still favors the full model over the minimal model (Fig 3C, Table 3C), with accuracies of 70±3% and 67±3% respectively (standard deviations calculated over 50 multiple imputations). The decrease in performance on the imputed data could be explained by the increase in error and variance due to multiple imputation, but the small number of patients with complete records in the KGH cohort left us with no better alternative for producing a more comprehensive evaluation, other than discarding the incomplete data altogether. The sensitivity and specificity values on both the complete and imputed data across a range of clinically-meaningful risk thresholds for outcome classification (Suppl. Table 1), show that the full model is consistently better at correctly predicting fatal cases, while misclassifying fewer non-fatal ones. Inspection of the calibration curves (Fig 3B and 3D) shows that both models systematically underestimate the observed risks. For example,

patients with a predicted mortality risk of 40% have an observed risk of over 60%, which is consistent with the fact that mortality among KGH patients (73%) is higher than in the training IMC cohort (58%).

External validation on the 158 EVD-positive patients in the GOAL dataset shows that the full and minimal models are able to accurately predict outcome for 71% and 69% of patients respectively (Fig 3E, Table 3D). In contrast to the KGH data, however, the models overestimate observed risk until approximately 60%, after which prediction estimates of the full model closely follow the observed risk (Fig 3F), which could also be seen as a consequence of the two datasets, IMC and GOAL, having similar CFRs (58% and 60%, respectively). We did not perform imputation on the GOAL dataset, because only a few cases (less than 10) had missing values.

Table 3: Model performance

Table 3A

	Full Model	*Minimal model*
Adjusted R^2	0.226 (0.179, 0.276)	0.174 (0.133, 0.217)
Apparent AUC	0.787 (0.743, 0.825)	0.771 (0.725, 0.811)
Corrected AUC	0.748 (0.697, 0.792)	0.759 (0.710, 0.801)
Brier score	0.208 (0.163, 0.255)	0.203 (0.159, 0.250)
Calibration	0.018 (0.010, 0.027	0.019 (0.011, 0.029)
Accuracy	0.687 (0.631, 0.736)	0.678 (0.622, 0.728)

Table 3B

	Full Model	*Minimal model*
AUC	0.98	0.78
Brier score	0.19	0.21
Calibration error	0.15	0.08

Accuracy	0.75	0.67
Sensitivity	0.70	0.62
Specificity	1.00	0.78

Table 3C

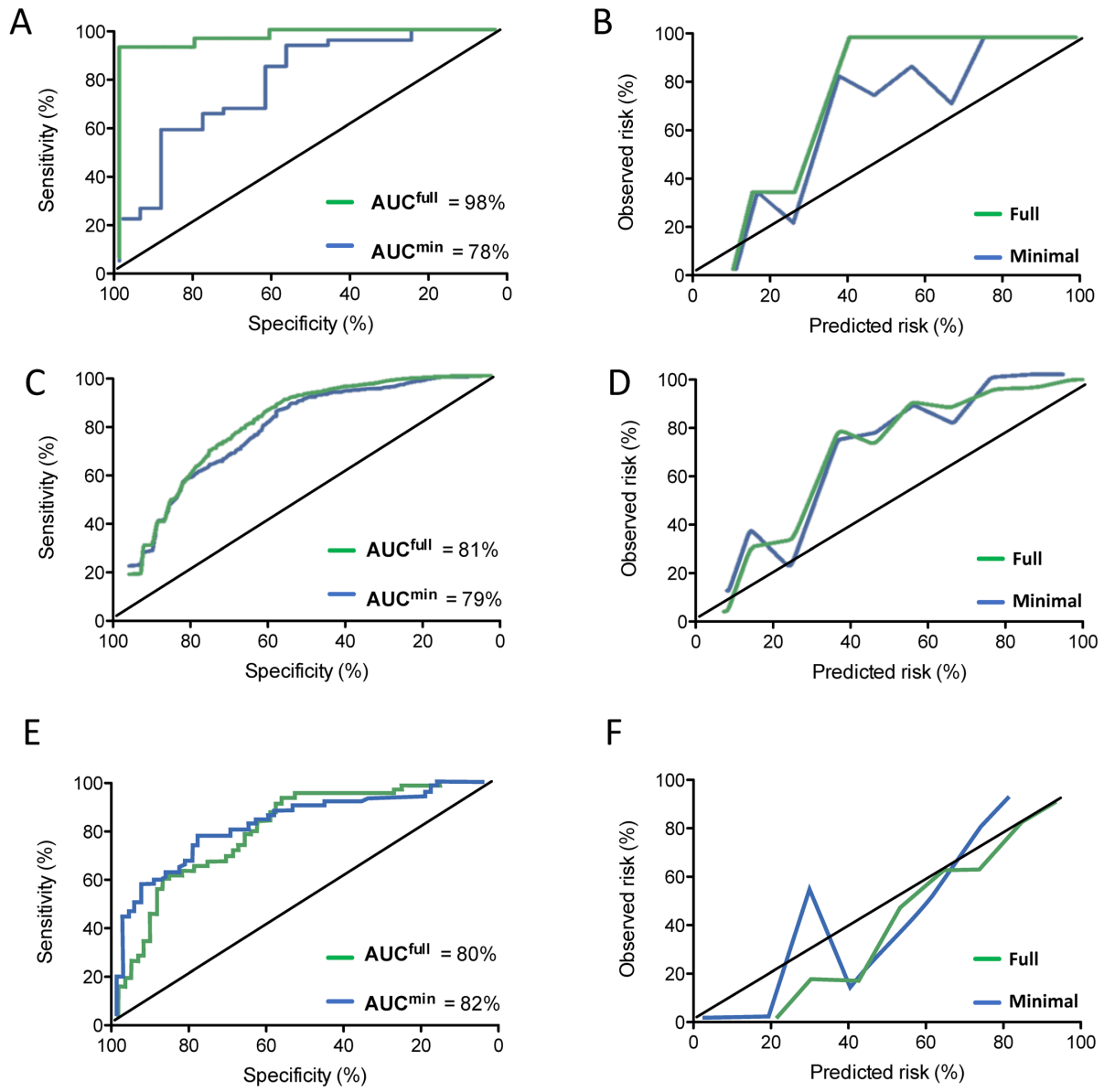
	Full Model	*Minimal model*
Corrected AUC	0.81 ± 0.03	0.79 ± 0.03
Brier score	0.20 ± 0.01	0.21 ± 0.01
Calibration error	0.07 ± 0.01	0.08 ± 0.01
Accuracy	0.70 ± 0.03	0.67 ± 0.03
Sensitivity	0.66 +/- 0.03	0.63 +/- 0.04
Specificity	0.82 +/- 0.06	0.80 +/- 0.05

Table 3D

	Full Model	*Minimal model*
AUC	0.80	0.82
Brier score	0.19	0.18
Calibration error	0.02	0.03
Accuracy	0.74	0.71
Sensitivity	0.95	0.92
Specificity	0.41	0.40

Table A shows the performance of the full and minimal models on the bootstrap samples taken from the IMC training set. Tables (B-D) report the performance of the minimal and full models during external validation, using either the complete Kenema General Hospital data (B), the imputed Kenema data (C) and the GOAL dataset (D). ***AUC:*** *Area under the Receiver-operating curve (ROC),* ***Brier score:*** *accuracy measure of assigning probabilistic predictions on mutually exclusive outcomes (ranging from 0 for perfect classification to 0.25 for a non-informative model with a 50% incidence of the outcome).*

Figure 3: Model performance in two independent external validations



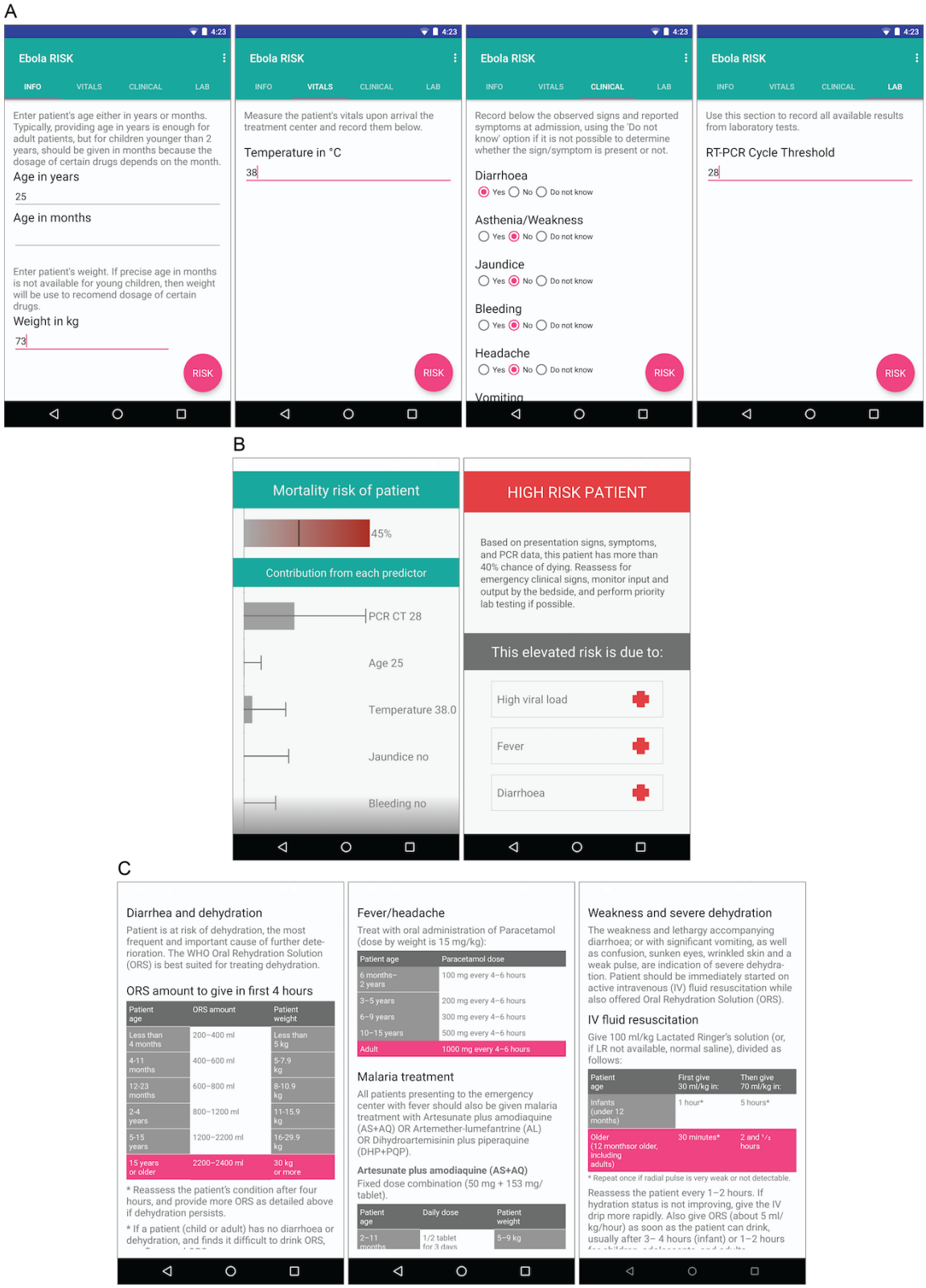
ROC (left) and calibration (right) curves as measures of performance for the full and minimal prognostic models. A-D show model performance for external validation using the Kenema General Hospital dataset, either when considering only complete records (A-B), or full data with missing values imputed (C, D). Graphs E-F show performance of a second external validation on

complete records from the GOAL dataset with 9% and 0% missing values for full and minimal models respectively.

4. Prognosis and risk visualization app

The app packages the full and minimal and models and it is available for free in the Google Play service named as "Ebola RISK." In order to allow iPhone users to test the app, we have also created an iOS version, distributed through Apple's App Store under the same name. The app selects the appropriate model according to the data provided by the user (Fig 4A),: it uses the full model if all required clinical sings/symptoms (body temperature, headache, bleeding, diarrhea, jaundice, vomit, abdominal pain, asthenia/weakness) at admission, patient age, and initial RT-PCR CT values are entered; otherwise, it defaults back to the minimal model, which only requires age and CT. Once the app selects a model based on the user input, it then estimates the mortality risk at admission, and stratifies the risk into three categories: low, medium and high. The high-risk threshold is set by default to 0.4, which brings the sensitivity very close to 1 for both the minimal and full models (as can be seen in Figures 2C and E, and in Suppl. Table 1), but at the expense of the specificity (although not dramatically: specificity goes under 40% only for the GOAL dataset when using a threshold lower than 0.5). The user can also choose this threshold and the desired balance of sensitivity vs specificity through the app's settings according to their clinical judgement, to better tailor their clinical triage based on resources and capacity. Furthermore, the detailed risk visualization with patient-specific charts (Fig 4B) can also be used to depict increases or decreases in a patient's risk due to changes in the signs and symptoms. This approach could help physicians to determine which factors would lead to the largest risk decreases, and to consider treatments that would be most effective to reduce risk of mortality.

Figure 4: Ebola RISK app



Data input, risk visualization, and clinical management screens in the Ebola RISK app. The data input screens (A) allow entering basic demographic information (age, weight), vitals, signs & symptoms at presentation, and lab results (CT value from first RT-PCR). Based on the available data, the app evaluates the death risk using either the minimal or full models, and presents a custom risk visualization (B). This visualization can either be a set of patient-specific charts (left screen in B), or a simplified list of the clinical features with a contribution to the risk score higher than a threshold (right screen in B). Selecting any of the features will open another screen with detailed information on the recommended treatment for that sign/symptom and adjusted according to the patient's age and weight (C).

The Ebola RISK app displays clinical protocols to treat and manage several conditions that often appear during the course of the illness, such as diarrhea, dehydration, fever, headache, and weakness (Fig 4C). The current version of the app does not include all of the protocols described in the WHO management manuals for patients with viral hemorrhagic fever, but only those that correspond to the predictors included in the prognostic models. The app is easily expandable to accommodate models with more predictors and additional clinical management information. The usage details of the app for both the Android and iOS versions are provided in the supplementary materials.

Discussion

The purpose of this study was two-fold: first, to present multivariate EVD prognostic models derived from the largest clinical multi-center dataset available to date, externally validated across diverse sites representing various periods of the epidemic; and second, to show how these models could guide clinical decisions by organizing existing knowledge of patient care and management more efficiently and making it easily available as a smartphone app. The robust performance of the models on the external sites indicate they could be generalized to new populations in future EVD outbreaks, a critical aspect for the app to be reliable and widely applicable. The IMC models recapitulate several findings reported earlier in the literature and also reveal further associations between mortality and clinical signs/symptoms. Occurrence of jaundice or bleeding at initial presentation is an important predictor of patient death, although both have a comparatively low incidence at triage among the patients in the IMC cohort of only 5%. More widespread EVD manifestations such as diarrhea and weakness have a much weaker correlation with mortality, at least based on their presence at triage, which seems to suggest that presentation of these clinical features says little about the clinical evolution of the patient.

The discriminative capacity of both full and minimal models is robust across the training set and the two independent testing sets, which were obtained at very distinct times during the epidemic, with AUCs ranging from 0.76 up to 0.82. The overall accuracy is consistently higher for the full model across the three sets, although the difference is no larger than 10%. While the most informative descriptors for predicting EVD outcome are patient age and viral load, more complex models offer higher accuracy by covering a larger proportion of the cohort. Inclusion of additional

predictors to the models, even those weakly associated with outcome, result in increased performance and improved stratification of observed patient outcomes. Our full model, incorporating several clinical signs/symptoms available at initial presentation – body temperature, jaundice, bleeding, weakness, headache, abdominal pain and vomiting – performs well on two independent datasets used for external validation differing by less than 3% AUC when values for the incomplete dataset were imputed. A major difference between these datasets was the time during which they were collected, with the KGH data representing an earlier time point, with less refined treatment protocols, higher viral virulence, increased patient volume and admission intensity with a larger number of patients delayed during transfers from holding centers. On the other hand, the GOAL dataset includes patients from the final months of the epidemic with a 13% lower CFR. Thus, as may be expected, the models underestimated the observed risk for patients of the KGH cohort, while observed risk was slightly overestimated in the GOAL cohort. The IMC training dataset covers a much broader temporal window of the epidemic as well as a wider catchment area, spanning several districts across two countries, which explain its robust performance in disparate populations.

Despite being the largest EVD prognosis modeling study to date, the amount and quality of available clinical data is still limited. We accounted for these limitations by applying various statistical techniques recommended for prognosis modeling (multiple imputation, bootstrap sampling, external validation), but ultimately future predictive models will require larger and better datasets. For example, in order to increase CT data, we aggregated measurements from different PCR labs, despite the use of different assays. Clinical signs/symptoms might be affected by variations in clinical assessments from a multitude of clinicians, and errors in the collection

(including patient symptom recall or history taking skills). Furthermore, clinical features are also limited as they often represent combinations of distinct clinical signs, which could result in certain features predicting a higher mortality alone but having the opposite effect in models that incorporate variables that better represent the causal relationships in the data. For instance, vomiting, headache, and abdominal pain have opposite correlation with mortality when controlled by age, viral load, and other factors in the full model. One plausible explanation could be that presentation of these symptoms prompts vigorous interventions from the health workers, such as oral rehydration, which have a significantly positive effect on outcome when the patient is not severely ill (as inferred from elevated viral load and young/advanced age.)

The prognosis and risk visualization app –while requiring further refinement, evaluation in the hands of its potential users, and ultimately validation through a clinical trial– shows the potential of data-driven medical apps to provide actionable clinical information based on specific patient characteristics. We are aware that, as a clinical decision support system designed with emergency treatment centers in low-resource areas in mind, our app faces several challenges, including potential for poor integration with clinical staff workflow, non-acceptance of computer recommendations, hardware failures, and insufficient training. We have considered some of these challenges by creating a minimal user interface that could be easily integrated into existing frameworks for field data collection, such as CommCareHQ and REDCap, and thus minimizing the need for separate training. We also simplified existing visualizations of patient risk derived from regression models, so that the predictions are displayed as clearly defined categories that could inform clinical intervention and resource management decisions. As new data would lead to more accurate and generalizable models, we designed the app so it can be updated regularly with

new models. The source code of the app is also available under the MIT open source license, to allow other researchers and developers to modify and extend the app without any restrictions.

In order to complement the risk prediction with actionable clinical information, the app displays the WHO recommended treatment guidelines applicable to each sign/symptom driving up the risk score. For instance, if the patient presents gastrointestinal symptoms such as vomiting and diarrhea, selecting them through the app's interface will open a new screen showing the WHO's rehydration protocol, customized by the age and weight of the patient. We sought to demonstrate how individual predictions from our prognostic model could assist health care workers in choosing the appropriate interventions based on patient's risk, physical characteristics, and observed manifestations of the disease. During the testing phase of the app, we noticed that a major fraction of the treatment protocols in the WHO guidelines (ORS administration, IV fluid resuscitation, blood transfusion, antibiotics, etc.) are explicitly indicated by the presentation of discrete symptoms. This observation suggests that the app could be used to provide a convenient and effective organization of the treatment strategies based on the symptoms entered by the clinician, with the added convenience of highlighting the appropriate doses and treatment recommendations according to the physical and demographic parameters of the patient. Although our current proof-of-concept app does not include the entirely of WHO's manuals for care and management of viral hemorrhagic fever patients (which encompass over 200 pages of clinical protocols), the provided information can be easily expanded in subsequent versions of Ebola RISK. This would make the app useful for personnel training and protocol adherence.

Finally, we aimed at developing a robust system that clinicians can trust in the field and in emergency situations. An initial step in that direction is to complement the prognosis predictions with authoritative clinical care information provided by sources such as the WHO. In this way, we envision the app both as a reference to improve training and adherence to protocol, as well as a support system that organizes clinical procedures more effectively around the patient's data. The integration of mHealth platforms with rapid point of care diagnostic kits (40, 41) has the potential to realize the concept of a “pocket lab” (42), which could be used outside laboratory settings and during health emergencies. However, success with such an integration can only occur after to properly addressing issues faced by these platforms in terms of validation and deployment (43), best practice standards (44), and regulatory oversight (14).

The performance of our models, and their reproducible predictive power with diverse external datasets, make us confident in our methods and reinforce our belief that the models could be useful in the near future to aid in stratification of patient support in the context of limited resources, as well as to serve as a benchmark for new models. In addition, they could be used in the design for Randomized Controlled Trials of EVD therapeutics or vaccines, where the models could serve as a standardized proxy for mortality. Fundamentally, the earlier we can make risk assessments, the earlier we can identify those in need of more intensive monitoring and treatment, which may not only improve outcomes but also better allocate limited resources. Clinicians and other health personnel making these assessments could support their clinical judgement with data-derived prognosis tools, such as the Ebola RISK app, as long these tools are based on accurate models that can be generalized to new patients. Our study indicates that the IMC-derived models are indeed generalizable. We believe that if clinical staff can obtain actionable information from these data-

derived tools, then they may be incentivized to generate more and higher-quality data. These data could be incorporated back into the models, thus creating a positive feedback loop. The use of low-cost tools on the ground, in combination with effective data collection and sharing among all stakeholders, will be key elements in the early detection and containment of future outbreaks of Ebola and other emerging infectious diseases.

Availability of source code, data, and app

The source code of all the modeling steps, from parameter fitting to internal and external validation, is openly available as a fully documented Jupyter notebook, deposited online at https://github.com/broadinstitute/ebola-imc-public. Refer to IMC's Ebola Response page (https://internationalmedicalcorps.org/ebola-response), for instructions on how external researchers can access the data. The prototype app is freely available on Google Play: https://play.google.com/store/apps/details?id=org.broadinstitute.ebolarisk, as well as on Apple's Play Store, under the Ebola RISK name.

Acknowledgments

We would like to thank the governments of Liberia, Sierra Leone, and Guinea for contributing to International Medical Corps' humanitarian response. We would also like to thank all of our generous institutional, corporate, foundation, and individual donors who placed their confidence and trust in International Medical Corps and made our work during the Ebola epidemic possible. We would also like to thank the United States Naval Medical Research Center, Public Health England, the European Union Mobile Laboratory, and the Nigerian Laboratory for providing laboratory data to our Ebola Treatment Units. We would like to further acknowledge all members of our Research Review Committee and other technical teams that contributed to this research. Finally, we would also like to thank our clinical, WASH, and psychosocial teams as well as all of our monitoring and evaluation staff, including the data collection officers at each of our ETUs. AC would like to thank Mary Lynn Baniecki and Christian Matranga for insightful discussions on the EBOV qPCR assays, and Christopher Moxon for critical feedback on the manuscript. AC and PCS are funded by the Howard Hughes Medical Institute and by NIH/NIAID 1R01AI114855.

References

1. Amarasingham R, Patel PC, Toto K, Nelson LL, Swanson TS, Moore BJ, et al. Allocating scarce resources in real-time to reduce heart failure readmissions: a prospective, controlled study. BMJ Quality & Safety. 2013;22(12):998-1005.

2. Mitchell M, Hedt-Gauthier BL, Msellemu D, Nkaka M, Lesh N. Using electronic technology to improve clinical care – results from a before-after cluster trial to evaluate assessment and classification of sick children according to Integrated Management of Childhood Illness (IMCI) protocol in Tanzania. BMC Medical Informatics and Decision Making. 2013;13(1):95.

3. Lund S, Boas I, Bedesa T, Fekede W, Nielsen H, Sørensen B. Association between the safe delivery app and quality of care and perinatal survival in ethiopia: A randomized clinical trial. JAMA Pediatrics. 2016;170(8):765-71.

4. Colubri A, Silver T, Fradet T, Retzepi K, Fry B, Sabeti P. Transforming Clinical Data into Actionable Prognosis Models: Machine-Learning Framework and Field-Deployable App to Predict Outcome of Ebola Patients. PLoS neglected tropical diseases. 2016;10(3):e0004549.

5. Schieffelin JS, Shaffer JG, Goba A, Gbakie M, Gire SK, Colubri A, et al. Clinical illness and outcomes in patients with Ebola in Sierra Leone. N Engl J Med. 2014;371(22):2092-100.

6. Hunt L, Gupta-Wright A, Simms V, Tamba F, Knott V, Tamba K, et al. Clinical presentation, biochemical, and haematological parameters and their association with outcome in patients with Ebola virus disease: an observational cohort study. The Lancet Infectious Diseases. 2015;15(11):1292-9.

7. Zhang X, Rong Y, Sun L, Liu L, Su H, Zhang J, et al. Prognostic Analysis of Patients with Ebola Virus Disease. PLoS neglected tropical diseases. 2015;9(9):e0004113.

8. Yan T, Mu J, Qin E, Wang Y, Liu L, Wu D, et al. Clinical characteristics of 154 patients suspected of having Ebola virus disease in the Ebola holding center of Jui Government Hospital in Sierra Leone during the 2014 Ebola outbreak. Eur J Clin Microbiol Infect Dis. 2015:1-7.

9. Li J, Duan H-J, Chen H-Y, Ji Y-J, Zhang X, Rong Y-H, et al. Age and Ebola viral load correlate with mortality and survival time in 288 Ebola virus disease patients. International Journal of Infectious Diseases. 2016;42:34-9.

10. Crowe SJ, Maenner MJ, Kuah S, Erickson BR, Coffee M, Knust B, et al. Prognostic Indicators for Ebola Patient Survival. Emerging Infectious Disease journal. 2016;22(2):217.

11. Steyerberg EW, Moons KGM, van der Windt DA, Hayden JA, Perel P, Schroter S, et al. Prognosis Research Strategy (PROGRESS) 3: Prognostic Model Research. PLOS Medicine. 2013;10(2):e1001381.

12. Goldstein BA, Navar AM, Pencina MJ, Ioannidis JP. Opportunities and challenges in developing risk prediction models with electronic health records data: a systematic review. Journal of the American Medical Informatics Association. 2016.

13. Hartley M-A, Young A, Tran A-M, Okoni-Williams HH, Suma M, Mancuso B, et al. Predicting Ebola Severity: A Clinical Prioritization Score for Ebola Virus Disease. PLoS neglected tropical diseases. 2017;11(2):e0005265.

14. Boulos MNK, Brewer AC, Karimkhani C, Buller DB, Dellavalle RP. Mobile medical and health apps: state of the art, concerns, regulatory control and certification. Online Journal of Public Health Informatics. 2014;5(3):229-.

15. Mosa ASM, Yoo I, Sheets L. A Systematic Review of Healthcare Applications for Smartphones. BMC Medical Informatics and Decision Making. 2012;12(1):67.

16. Sacks JA, Zehe E, Redick C, Bah A, Cowger K, Camara M, et al. Introduction of Mobile Health Tools to Support Ebola Surveillance and Contact Tracing in Guinea. Global Health: Science and Practice. 2015;3(4):646-59.

17. Balakrishnan R, Gopichandran V, Chaturvedi S, Chatterjee R, Mahapatra T, Chaudhuri I. Continuum of Care Services for Maternal and Child Health using mobile technology – a health system strengthening strategy in low and middle income countries. BMC Medical Informatics and Decision Making. 2016;16(1):84.

18. Kaphle S, Chaturvedi S, Chaudhuri I, Krishnan R, Lesh N. Adoption and Usage of mHealth Technology on Quality and Experience of Care Provided by Frontline Workers: Observations From Rural India. JMIR mHealth uHealth. 2015;3(2):e61.

19. Cherutich P, Golden M, Betz B, Wamuti B, Ng'ang'a A, Maingi P, et al. Surveillance of HIV assisted partner services using routine health information systems in Kenya. BMC Medical Informatics and Decision Making. 2016;16(1):97.

20. Steiner A, Hella J, Grüninger S, Mhalu G, Mhimbira F, Cercamondi CI, et al. Managing research and surveillance projects in real-time with a novel open-source management tool designed for under-resourced countries. Journal of the American Medical Informatics Association. 2016;23(5):916-23.

21. Tom-Aba D, Olaleye A, Olayinka AT, Nguku P, Waziri N, Adewuyi P, et al. Innovative Technological Approach to Ebola Virus Disease Outbreak Response in Nigeria Using the Open Data Kit and Form Hub Technology. PloS one. 2015;10(6):e0131000.

22. Sivakumaran M, Iacopino P. The Mobile Economy 2018. GSMA Intelligence; 2018.

23. WHO. Manual for the care and management of patients in Ebola care units / Community Care Centres: interim emergency guidance, January 2015. 2015.

24. WHO. Clinical management of patients with viral haemorrhagic fever. A pocket guide for front-line health workers. 2016.

25. Roshania R, Mallow M, Dunbar N, Mansary D, Shetty P, Lyon T, et al. Successful Implementation of a Multicountry Clinical Surveillance and Data Collection System for Ebola Virus Disease in West Africa: Findings and Lessons Learned. Global Health: Science and Practice. 2016;4(3):394-409.

26. Harrell FE. Regression modeling strategies : with applications to linear models, logistic and ordinal regression, and survival analysis. Second edition. ed. Cham: Springer; 2015. 1 online resource (xxv, 582 pages) p.

27. Frank E. Harrell J. Regression Modeling Strategies: Springer; 2015.

28. Lado M, Walker NF, Baker P, Haroon S, Brown CS, Youkee D, et al. Clinical features of patients isolated for suspected Ebola virus disease at Connaught Hospital, Freetown, Sierra Leone: a retrospective cohort study. The Lancet Infectious Diseases.

29. Miles A. Obtaining Predictions from Models Fit to Multiply Imputed Data. Sociological Methods & Research. 2015.

30. Efron B, Tibshirani R. Improvements on Cross-Validation: The 632+ Bootstrap Method. Journal of the American Statistical Association. 1997;92(438):548-60.

31. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the Performance of Prediction Models: A Framework for Traditional and Novel Measures. Epidemiology. 2010;21(1):128-38.

32. Yaniv I, Yates JF, Smith JK. Measures of discrimination skill in probabilistic judgment. Psychological Bulletin. 1991;110(3):611.

33. Harel O. The estimation of R 2 and adjusted R 2 in incomplete data sets using multiple imputation. Journal of Applied Statistics. 2009;36(10):1109-18.

34. Hartley M-A, Young A, Tran A-M, Okoni-Williams HH, Suma M, Mancuso B, et al. Predicting Ebola infection: A malaria-sensitive triage score for Ebola virus disease. PLoS neglected tropical diseases. 2017;11(2):e0005356.

35. Gire SK, Goba A, Andersen KG, Sealfon RSG, Park DJ, Kanneh L, et al. Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. Science (New York, NY). 2014;345(6202):1369-72.

36. Center USNMR. Ebola Zaire (EZ1) rRT-PCR (TaqMan) Assay. 2014.

37. Trombley AR, Wachter L, Garrison J, Buckley-Beason VA, Jahrling J, Hensley LE, et al. Comprehensive Panel of Real-Time TaqMan™ Polymerase Chain Reaction Assays for Detection and Absolute Quantification of Filoviruses, Arenaviruses, and New World Hantaviruses. The American journal of tropical medicine and hygiene. 2010;82(5):954-60.

38. Prevention CfDCa. Ebola Virus VP40 Real-Time RT-PCR Assay. 2014.

39. Van Belle V, Van Calster B. Visualizing Risk Prediction Models. PloS one. 2015;10(7):e0132614.

40. Broadhurst MJ, Kelly JD, Miller A, Semper A, Bailey D, Groppelli E, et al. ReEBOV Antigen Rapid Test kit for point-of-care and laboratory-based testing for Ebola virus disease: a field validation study. The Lancet.386(9996):867-74.

41. Yen C-W, de Puig H, Tam JO, Gomez-Marquez J, Bosch I, Hamad-Schifferli K, et al. Multicolored silver nanoparticles for multiplexed disease diagnostics: distinguishing dengue, yellow fever, and Ebola viruses. Lab on a Chip. 2015;15(7):1638-41.

42. Radin JM, Topol EJ, Andersen KG, Steinhubl SR. A laboratory in your pocket. The Lancet.388(10054):1875.

43. Tomlinson M, Rotheram-Borus MJ, Swartz L, Tsai AC. Scaling Up mHealth: Where Is the Evidence? PLOS Medicine. 2013;10(2):e1001382.

44. Misra S, Lewis T, Aungst T. Medical application use and the need for further research and assessment for clinical practice: Creation and integration of standards for best practice to alleviate poor application design. JAMA Dermatology. 2013;149(6):661-2.