Beyond Poverty as a Proxy

Reducing Inequality in Infant Mortality by Identifying and Targeting Higher Risk Births

Chad Hazlett, PhD CHAZLETT@UCLA.EDU
Department of Statistics
Department of Political Science
University of California, Los Angeles

Antonio P. Ramos, PhD TOMRAMOS@G.UCLA.EDU
Department of Biostatistics
Fielding School of Public Health
California Center for Population Research
University of California, Los Angeles

Stephen Smith, PhD Candidate STEPHENSMITH13424@GMAIL.COM
Department of Statistics
University of California, Los Angeles

Corresponding Author: Antonio P. Ramos, TOMRAMOS@G.UCLA.EDU

Research in Context

Evidence before the study
In many low-income countries, interventions that aim to reduce infant mortality cannot be made universally available. While low-cost medical interventions are available to avoid many of the most common causes of death, these interventions can only save lives if they are targeted to individuals who would have otherwise died. Currently, most national programs and policies aiming to reduce infant mortality employ only poverty-related assessments to target those programs. Large gains in the effective targeting of such programs, and thus their life-saving potential, may be possible if risk can be assessed using more flexible statistical models that draw on a variety of easy-to-collect data, but existing models have not. Surprisingly, few prior studies have combined multiple risk factors to estimate early-life mortality risk in the general population.[1,2,3] None of these employ flexible (machine learning) models, or compare the results to a traditional poverty- or wealth-based targeting benchmark.

Added value of this study
Working with publicly available data from 22 countries in sub-Saharan Africa, we first show that targeting early mortality interventions based only on income assessments cannot do much better than random distribution in most countries. We then employ established machine learning methods to show how using a range of generally available data prior to each birth, a risk score can be computed for each birth that is more predictive of early mortality. Targeting the highest risk decile of births, for example, can capture only about 10% of deaths using income alone, but 15-30% (depending on the country) with our more inclusive model.

Implications of all available evidence
We conclude that it is feasible, using standard machine learning methods and data similarly to what is routinely collected, to greatly improve upon income-based targeting schemes for interventions seeking to minimize early mortality. In the cases studied, these could greatly improve the targeting and consequently the life-saving potential of these interventions.

Abstract

Background: Infant mortality remains high and unevenly distributed in much of sub-Saharan Africa. Though cheap and effective interventions can in principle prevent many of these deaths, given a finite budget for supplying such interventions in many countries, their life-saving benefit is critically constrained by the ability to target them to those

individuals who would have otherwise have likely died. While countries routinely use poverty- or wealth-based thresholds to target policies and interventions including those related to maternal health and early mortality, we examine whether models that consider other easy-to-measure and factors can substantially improve our ability to identify higher-risk births, and consequently the potential for life-saving intervention.

Methods: Using up to 25 variables from the demographic health survey for 22 sub-Saharan African countries, we employ established machine learning methods to estimate per-child risk of infant mortality, separately in each country. We use out-of-sample testing methods to provide an honest assessment of how useful these methods are for identifying high risk births. We also compare this to a benchmark similar to standard practice, in which only (within-country) measures of wealth can be used for targeting.

Results: Targeting based on wealth alone is only slightly better than random targeting: the poorest 10 percent of the population experiencing approximately 10 percent of total infant mortality burden. By contrast the 10 percent of the population at highest risk according to our model account for 15-30 percent of infants deaths, depending on country. In other words, for a hypothetical intervention with some fixed effectiveness and that can be administered to only a chosen 10 percent of the population, such a targeting approach would save 1.5-3 times as many lives as one based on wealth alone.

Interpretation: In the 22 sub-Saharan African countries studied, approaches that use commonly available types of data to flexibly predict high-risk births has the potential to substantially improve the targeting and thus maximum potential effectiveness of life-saving interventions against early mortality.

Goal 3.2 of the Sustainable Development Goals (SDG) seeks to "end preventable deaths of newborns and children under 5 years of age, with all countries aiming to reduce neonatal mortality to at least as low as 12 per 1,000 live births and under-5 mortality to at least as low as 25 per 1,000 live births." Despite large overall improvements, including a 44% reduction in child mortality globally from 2000 to 2015, progress has varied widely both between and within countries,[4,5,6] with under-5 mortality exceeding 80 per 1,000 live births in some countries of sub-Saharan Africa. Further, improvements in national averages can mask large and widening gaps between sub-populations, with more privileged groups improving their health status faster than others.[7,8]

Many of these deaths are considered preventable as they can very effectively be avoided with existing, low-cost interventions such as oral rehydration therapy, malaria prevention and treatment, vaccines (e.g. against measles), and nutritional support.[9,10] A major limitation, however, is that most countries with high early mortality cannot provide such interventions universally, so must choose how they should be targeted. The accuracy of this targeting severely limits the life-saving potential of an intervention, because an intervention given to a child who would have survived without that intervention has no

effect on mortality. An infant mortality rate of 80 deaths per 1,000 births for example is far too high, yet this rate of 0.8% means that intervention given to a random set of individuals can save at most 0.8% of those to whom it is given. More realistically, such an intervention will only have some efficacy rate among those who would have otherwise have died, say, 10%.[1] An intervention given at random could then save the lives of only 0.8% * 10% = 0.08% of those to whom it is given. By contrast, if such an intervention could instead be targeted to a group with a higher risk of mortality, say 10%, then with the same efficacy rate it would save the lives of 10% * 10% = 1% of recipients, saving 12.5 times as many lives. In this way, the lives saved by a (hypothetical) intervention with a given efficacy rate is limited by, and proportional to, the ability to target those at risk. Efforts to improve targeting should thus be treated with as much importance and urgency as efforts to improve the efficacy of the available intervention.

Unfortunately, this has not been the case. Although early mortality can have many causes and can be related to many risk factors, most countries facing high early mortality target policies and interventions based on one or a few risk factors, almost invariably indicators for low income or wealth, or a related conception of poverty. But research into models for early mortality that look beyond income, poverty, or wealth remain rare, and not designed to demonstrate how practical, national level systems could target interventions to high-risk births. A recent, partial exception[11] finds, first, that in Brazil and India, poverty is a poor proxy for risk of early mortality, much as we find below for countries in sub-Saharan Africa. Further, for births in India, they find that using just four risk factors in a linear model roughly doubles the fraction of infant deaths that can be identified in the highest risk quintile as compared to models relying on a proxy for poverty. Our work builds on this by examining the degree to which flexibly-estimated, individual-level risk scores based only on a range of generally available, pre-birth data can improve the targeting of life-saving interventions against infant mortality. Specifically, we estimate mortality risk in individual infants (under one year of age) in 22 countries in sub-Saharan Africa. We employ flexible (machine learning) models able to identify predictive relationships that may have eluded more rigid regression models, particularly as we provide these models with access to a wider range of risk factors than previous work. We also take care to employ only risk factors that can generally be known or anticipated prior to a child's birth, as practical targeting systems may be limited to such data. As is standard in many machine learning applications, we employ cross-validation to protect against over-fitting overly optimistic performance claims. However, breaking with convention, we rely on mathematically simple, easy-to-understand performance metrics that directly relate to the potential impact of a hypothetical policy intervention.

To understand how this compares to the status quo, we benchmark our models against a wealth-based model. This is an imperfect proxy for the policies that in fact use poverty-line determinations to target (bundled) programs, as poverty assessment may be in part or whole based on income rather than wealth. That said, at least some programs that target poor household use data sources to determine poverty status very similar to ours.[12]

[1] By efficacy rate, we mean the fraction of people who would have otherwise died, but who instead survive when given this intervention. This is sometimes known as the probability of necessity and sufficiency XXX CITE.

Further, our wealth-based approach may be conservative in that it could over-estimate the efficacy of such approaches because (i) wealth, not just income, is likely to play a role in early mortality, and (ii) our models flexibly use a within-country wealth index to estimate risk of early mortality, without assuming a single binary cutoff or any other particular functional form for the relationship.

Materials and Methods

Data Sources

We employ the most recent available Demographic and Health Surveys (DHS), as augmented and prepared for analysis by the Integrated Public Use Microdata Series (IPUMS), henceforth referred to as the IPUMS DHS data. Like other work studying early mortality using the IPUMS DHS data (e.g. [13,14]), we use the "child module" which contains data on births for children born up to five years prior to the interview of their mother. While this restricts our attention to children under the age of five, this module contains more predictive variables than the "all births" module. More importantly, it would be inadvisable to use data collected on the mother, household, malaria prevalence, location, etc. in the year of the DHS data collection to predict early mortality from a birth many years before.

Originally, the IPUMS DHS data contain surveys on 28 sub-Saharan African countries. We consider geographic information to be critical for these models, since (i) even coarse geographic indicators can be highly predictive of risk and reduce the need for other, harder to measure variables associated with location, (ii) we considered malaria prevalence a potentially important variable, but because it is computed for a 10km area around the DHS sample cluster centroid, it is only available in surveys that collected that contain those geographic coordinates, and (iii) the ability to target interventions partly by geography is convenient for policymakers. Unfortunately in six countries (Burundi, Chad, Ethiopia, Liberia, Namibia, Sudan) the IPUMS DHS data did not include geographic indicators even at the district level. This limits us at present to surveys from 22 countries: Angola (2015), Benin (2011), Burkina Faso (2010), Cameroon 2011, Democratic Republic of Congo (2013), Cote d'Ivoire (2011), Ghana (2014), Guinea (2012), Kenya (2014), Lesotho (2014), Madagascar (2008), Malawi (2016), Mali (2012), Mozambique (2011), Niger (2012), Nigeria (2013), Rwanda (2014), Senegal (2017), Tanzania (2015), Uganda (2016), Zambia (2013), Zimbabwe (2015). Table 1 provides additional information on the available data within each of these countries.

Variables

Our outcome is infant mortality, indicated by death before reaching one year of age.[2] For predictors or risk factors (we use the terms interchangeably) we limit ourselves to

[2]Mortality before age five (under-five mortality) is also commonly of interest, but requires waiting until all children in the dataset reach age 5, which would eliminate most of the data in our dataset.

predictors that (i) would be reasonably available to policymakers and health workers in-country, and (ii) are measurable prior to the birth of the child, rather than depending on the health or other features of the child once borne. The latter constraint is in place so that these risk factors could be used to calculate risk in advance and deploy interventions to households, clinics, or regions with high risk.

The resulting risk factors we consider are: maternal age, malaria prevalence in the vicinity, head of household age and sex, parental education, cooking fuel type, floor type, toilet type, drinking water type, income, birth order, birth month, religion, bed net usage, age at first marriage or union, and the death of a previous child under one or five years of age (see Table S3 in the online supplement). Geographic information is also used, but varies in its granularity by country. We recode variables with many categories to construct measures that are likely to perform better. For example, the DHS data initially contain 66 categories of toilet types, which we simplified to 4 categories based on the level of improvement. Table 1 indicates, for each country, which of these variables (if any) is not available in the DHS data, the number of observations (original and with complete data), the number of infant deaths, and a description of the type of geographic data available. Note that, because the machine learning methods we apply can tolerate a large number of variables and choose which to ultimately included "on their own", we do not perform a separate model selection procedure to narrow this set of variables.

Finally, the IPUMS DHS data do not contain an indication of how an individual or household is classified relative to the poverty line in a given country, nor does it attempt to measure income as such. Thus, rather than including any locally-defined poverty measure, we employ the wealth-based indices provided by DHS. This measure is an imputed index for within-country wealth, imputed by DHS using principal components analysis over a number of variables including ownership of radios, televisions, and other domestic equipment; electricity and clean water; type of materials used in the walls, floor and roof; and the type of toilet in the household. Scores are calculated at the household level. The only requirement we make of this measure is that, within a given country, it provides an imperfect but potentially informative measure of relative wealth. Since this index is constructed within each country, and we train each country's risk model separately, there is no expectation nor need for cross-country comparability on this index.

Modeling

We estimate infant mortality risk for each birth in the data set. In each country, we separately run two sets of models. One is a "wealth model", described below as a proxy for what can be achieved through poverty-based targeting. The other set contains flexible machine learning approaches that use the predictors described above to estimate individual level mortality risk. These include random forest (rf), extreme gradient boosting

Similarly, studying under-5 mortality also requires limiting ourselves to older data that are less relevant to the present or to future years in which these models would be applied.

Table 1: Data availability by country

Country	N original	N complete	N deaths	Geography	Unavailable
Angola	11024	10829	472	province (18)	age first marriage, malaria
Benin	10490	10199	471	department (12)	
Burkina Faso	11728	10905	752	region (13)	
Cameroon	9041	8450	564	region or city (12)	
Congo (D.R.)	14384	12731	785	province (11)	previous child death(s)
Cote d'Ivoire	5945	5211	421	none	
Ghana	4548	4203	189	none	
Guinea	5422	5223	396	region (8)	cooking fuel
Kenya	16557	15110	589	none	pregtermin
Lesotho	2360	2127	145	district (10)	bednets, malaria
Madagascar	9657	9229	459	region (22)	cooking fuel
Malawi	13633	13231	563	region (3)	malaria
Mali	8043	7913	475	region (6)	cooking fuel, malaria
Mozambique	8493	8026	537	province (11)	
Niger	9650	9612	586	region (8)	religion, malaria
Nigeria	24222	23559	1686	geopolitical zone (6)	
Rwanda	6055	5350	168	province (5)	cooking fuel
Senegal	9585	9371	401	region (14)	malaria
Tanzania	8007	7688	319	district (10)	cooking fuel, religion, malaria
Uganda	12089	11619	484	district (112)	cooking fuel, malaria
Zambia	10646	9811	483	none	
Zimbabwe	4852	4630	236	province (10)	malaria

Note: Summary of data by country. The DHS data initially include `N original` observations in each country, which falls to `N complete` when we remove missing values on the included variables. `Geography` indicates what type of geographic information, if any, is available and the number of units in parentheses. All other variables described in text are used in each country except where otherwise indicated under `Unavailable`.

over trees (xgb), kernel regularized least squares (krls), and elastic-net regularized logistic regression (elastic-logit).

Both rf[15] and xgb[16] are tree-based approaches. The simplest tree-based approaches are classification and regression trees (cart[17]), which recursively partition observations into groups based one cutoff values of the predictor variables, so that each group has similar values of the outcome. While such single-tree models are easy to interpret, they often perform poorly, and small changes in the input data can have dramatic effects on the model and results. The rf and xgb methods improve on single-tree approaches by combining many trees. For rf, this is achieved by effectively averaging together the estimates of many trees, each employing only a random subset of the input variables and trained on a random subset of the observations. As applied here, each forest contains 500 trees, and the number of variables randomly sampled for each tree split (`mtry`) was set for each country's model using cross-validation. Under xgb, this is similarly achieved by combining many trees, but by beginning with one tree, adding the predictions of another tree trained on the errors of the prior, then repeating this process hundreds or thousands of times. We use cross-validation for each country to determine (i) the learning rate ($\eta \in (0, 1)$, where smaller values help to prevent overfitting), (ii) the maximum depth of the trees (`max_depth`), (iii) the proportion of the variables to be considered for tree construction (`colsample_bytree`), (iv) the proportion of observations from the training set used for modeling (`subsample`), and (v) the number of iterations for the boosting procedure (`nrounds`). The remaining parameters were set at common default values: the minimum requirement of prediction improvement before selecting a more complex model

(γ) was set to 0, and `min_child_weight`, where higher values restrict the depth of each tree based on a measure of the homogeneity of the labels within the nodes (a sign of overfitting), was set to 1.

Kernel-based regularized least squares (krls[18]) is a regression and classification technique that employs kernels. Such approaches effectively build a model by leveraging information about how similar observations are to each other. Each observation is treated as a potential "fence post" in the data, and the model learns how being closer to or further from each such location in the data influences the expected probability of mortality. We employ a Gaussian kernel as the measure of similarity using default values.[18]

Elastic-net logistic regression (elastic-logit[19]) is similar to conventional logistic regression. While ordinary logistic regression chooses coefficients β_j on each variable j so as to maximize the likelihood by minimizing the negative log-likelihood, elastic-net logistic regression minimizes the negative log-likelihood *plus* a penalty term that helps to regularize or "shrink" the model, preventing overfitting. In our case that penalty term is equal to $\lambda(\frac{\alpha}{2}\sum_j |\beta_j| + \frac{1-\alpha}{2}\sum_j \beta_j^2)$. The parameters α and λ are chosen by cross-validation. This constitutes a compromise between what is known as ridge regression (which is penalized by the ℓ_2 norm, $\sum_j \beta_j^2$), and LASSO (penalized by the ℓ_1 norm, $\sum_j |\beta_j|$). By including the ℓ_1 norm in the penalty, elastic-net has the benefit of setting some coefficients to exactly 0, that is, selecting some variables while dropping others.

Finally, we construct a simple ensemble model in which each predicted probability is the unweighted average of the probabilities produced by elastic-logit, rf, xgb, and krls. This has two primary benefits. The first is that such ensembles often have performance as good as the best sub-models and sometimes better. Second, it allows us to commit to reporting results for a single model—the ensemble model—as our premier estimate of the predictive power of these model on the test data. This is preferable to reporting the best-performing model on the test set, which can be misleading because such a choice is effectively a means of "training on the test set".

Wealth model

Our wealth-based model employs wealth information flexibly: mortality is regressed on ten indicators for decile of wealth, equivalent to finding the mean mortality level in each wealth decide. In evaluating model performance, we can consider how being in the highest risk decile, quartile, or other groups is related to mortality risk. There is no presumption that risk is necessarily decreasing or even monotonic in the level of wealth.

We note that this is a wealth model, not a poverty model based on poverty assessments made by countries, usually based on income. Neither income nor poverty-assessments themselves are available in the DHS data. This limits our ability to benchmark against a true poverty-based system. That said, our wealth model may reasonably provide an optimistic indication of what can be predicted by income or poverty-based approaches. Household income is time-sensitive, but is likely to generate risks of early mortality when it remains low on average over a long period of time, in which case wealth will also drop.

Moreover, financial shocks such as large expenses or losses may have little impact on income, but substantial impact on quality of life, access to healthcare, and poverty as locally understood (see e.g.[20]). Our wealth model is also more flexible and less coarse than poverty-based targeting systems, which typically assign individuals through a binary system (e.g. above or below the poverty line). Again this is intended to be a favorable choice for the wealth model, so that we can later consider its demonstrated performance likely to be over- rather than under-stated.

Sample splitting and cross-validation

Before fitting any model, we split the data into training and test sets. For each country, 80% of observations are placed in a training and validation set, while 20% are retained in a held-out testing set. For the models described above, we tune hyperparameters using either 10-fold cross-validation within the training set (for elastic-logit, rf, xgb) or leave-one-out cross-validation (for krls). The cross-validation procedure also gives a preliminary assessment of model performance. Once all modeling choices were finalized based on this cross-validation, those models were "frozen" and we then assessed performance on the test set. This was done only once, after writing the a draft of the paper using the cross-validation results, then updating the final values using the actual test-set results. This provides an honest out-of-sample performance assessment, prohibiting any "training on the test set". We report the accuracy of the models using the performance metrics described below.

Variable importance

Our models only seek to determine the best risk estimate possible given a set of predictors and the answers say nothing of the causal effect of those predictors on the risk of mortality. Indeed, many of the risk factors, such as religion or the type of roof, are valuable not because they cause mortality but because they are expected to signal the presence of other, unobserved factors that cause higher mortality. In other words, confounding is expected and hoped for in these models.

It is valuable nevertheless to understand which variables the models deem most important in estimating the risk, for the purposes of assessing the credibility of the model, improving future models, or focusing public health resources on data collection. Unfortunately, for many machine learning models it is more difficult to characterize the model's behavior than is the case with simpler regression approaches like OLS or logit, in which the model is fully characterized by a small number of parameters. That said, readily available measures of variable importance for different models can provide some useful insights. Specifically, for elastic-net we report how often the model chose a given variable and the proportion of time their coefficient is positive. For random forest, the (scaled) "variable importance" describes what fraction of the time a given variable was chosen to be included in the classification trees over which each random forest model aggregates.

Performance metrics

Finally, to evaluate our approach, we employ simple, policy-relevant metrics rather than more esoteric and difficult to interpret metrics often relied upon in machine learning tasks. A model's *recall* (also referred to as sensitivity) is the proportion of actual deaths that are included in the group we consider to be at risk. What we term "recall10" is defined by first determining the group of births in the top risk decile (hence the "10") according to our model, then computing the fraction of all actual infant deaths accounted for by this group. In policy terms, this corresponds to a case in which we have resources to target 10% of the population, and wish to know how much of the total mortality risk would be covered in that group. An ineffective risk assessment that assigns risk at random would produce a recall10 of 0.10 (10%). By contrast an effective system points us to a top risk decile that accounts for more than 10% of the mortality, such as 20% or 30%. Similarly, we report recall20, which tells us the fraction of early deaths that occurred within the top risk quintile. For both recall10 and recall20, we also report "efficiency gain", which tells us how many times more effective the ensemble model is than the benchmark wealth model.

Results

Poor performance of the wealth model. Table 2 provides the primary results, showing the proportion of mortality that occurred among the top 10% highest predicted risk individuals in the test set. In Rwanda and Uganda, wealth is a reasonable indicator of risk, and the 10% of the population with the higher wealth-based predicted risk account for 24% and 18% of the mortality, respectively. In the other 20 countries, wealth is almost uninformative, with recall10 values falling between just 3% (Tanzania) and 14% (Lesotho). Overall, the (country-wise) average recall10 of just 11% suggests that the wealth generally model does little or no better than chance in predicting which individuals are at highest risk. Similarly, Table 3 shows that in most countries, recall20 is only slightly above 20%, averaging (country-wise) to 23%. Rwanda is again an exception, with the top quintile accounting for 48% of mortality, while recall20 rates range from 11% to 29% in the others, and are at or below chance (20%) in 9 of the 22 countries.[3]

Beyond wealth: Gains from richer models. The elastic-net logit, rf, xgb, and krls models all augment the wealth model by adding the additional variables described above. Overall these models perform similarly to each other, and far better than the wealth-only model. The top risk deciles according to these models account for, on average, 17-20% of mortality (recall10); the top risk quintile accounts for 30-32% of mortality (recall20). The ensemble model performs better than any of the individual models, on average capturing

[3]Appendix A provides the same performance metrics at Tables 2 and 3, but using cross-validated results from the training set. Those results are extremely similar to the test set results reported here, indicating very little over-fitting in the training process.

Table 2: recall10 results, test set

Survey	wealth	elastic-logit	rf	xgb	krls	ensemble	efficiency gain
Angola 2015	0.11	0.23	0.23	0.23	0.22	0.27	2.50
Benin 2011	0.13	0.22	0.31	0.20	0.21	0.22	1.75
Burkina Faso 2010	0.12	0.15	0.21	0.17	0.15	0.19	1.61
Cameroon 2011	0.10	0.18	0.23	0.17	0.17	0.24	2.45
Congo (D.R.) 2013-14	0.13	0.14	0.18	0.19	0.14	0.18	1.40
Cote d'Ivoire 2011	0.12	0.21	0.17	0.27	0.20	0.24	2.00
Ghana 2014	0.05	0.11	0.19	0.16	0.08	0.19	3.50
Guinea 2012	0.11	0.14	0.20	0.16	0.15	0.19	1.67
Kenya 2014	0.06	0.21	0.17	0.12	0.19	0.21	3.57
Lesotho 2014	0.14	0.14	0.14	0.14	0.10	0.17	1.25
Madagascar 2008	0.08	0.10	0.15	0.10	0.13	0.21	2.71
Malawi 2016	0.12	0.12	0.21	0.20	0.12	0.19	1.62
Mali 2012	0.07	0.20	0.21	0.23	0.23	0.23	3.14
Mozambique 2011	0.09	0.13	0.21	0.13	0.13	0.21	2.20
Niger 2012	0.12	0.15	0.13	0.17	0.13	0.16	1.36
Nigeria 2013	0.12	0.21	0.22	0.19	0.24	0.25	2.15
Rwanda 2014	0.24	0.21	0.09	0.21	0.21	0.21	0.88
Senegal 2017	0.10	0.16	0.26	0.29	0.15	0.29	2.87
Tanzania 2015	0.03	0.21	0.14	0.16	0.17	0.22	7.00
Uganda 2016	0.18	0.14	0.24	0.24	0.25	0.26	1.47
Zambia 2013	0.10	0.20	0.23	0.16	0.19	0.21	2.00
Zimbabwe 2015	0.13	0.13	0.21	0.19	0.17	0.19	1.50
Average	0.11	0.17	0.20	0.19	0.17	0.22	2.30

Table 3: Recall20 results, test set

Survey	wealth	elastic-logit	rf	xgb	krls	ensemble	efficiency gain
Angola 2015	0.27	0.45	0.34	0.38	0.43	0.44	1.64
Benin 2011	0.28	0.36	0.46	0.34	0.39	0.46	1.65
Burkina Faso 2010	0.26	0.35	0.26	0.33	0.41	0.36	1.38
Cameroon 2011	0.29	0.27	0.33	0.29	0.29	0.34	1.19
Congo (D.R.) 2013-14	0.20	0.26	0.32	0.29	0.29	0.31	1.55
Cote d'Ivoire 2011	0.25	0.44	0.26	0.40	0.36	0.43	1.71
Ghana 2014	0.16	0.16	0.22	0.27	0.16	0.24	1.50
Guinea 2012	0.18	0.27	0.34	0.24	0.33	0.33	1.86
Kenya 2014	0.17	0.26	0.32	0.25	0.31	0.32	1.90
Lesotho 2014	0.24	0.28	0.28	0.24	0.14	0.24	1.00
Madagascar 2008	0.16	0.26	0.30	0.24	0.25	0.31	1.87
Malawi 2016	0.26	0.23	0.30	0.31	0.25	0.30	1.17
Mali 2012	0.17	0.38	0.35	0.36	0.37	0.39	2.31
Mozambique 2011	0.21	0.23	0.32	0.22	0.29	0.34	1.64
Niger 2012	0.21	0.27	0.21	0.29	0.24	0.28	1.38
Nigeria 2013	0.22	0.39	0.35	0.37	0.40	0.40	1.81
Rwanda 2014	0.48	0.30	0.33	0.39	0.36	0.33	0.69
Senegal 2017	0.19	0.30	0.38	0.39	0.30	0.40	2.13
Tanzania 2015	0.11	0.32	0.29	0.35	0.25	0.29	2.57
Uganda 2016	0.25	0.27	0.33	0.39	0.34	0.42	1.67
Zambia 2013	0.26	0.32	0.34	0.31	0.32	0.31	1.20
Zimbabwe 2015	0.19	0.26	0.26	0.32	0.21	0.36	1.89
Average	0.23	0.30	0.31	0.32	0.30	0.35	1.62

22% of the mortality in its top risk decile (recall10), and 35% in the top risk quintile (recall20). Compared to the wealth model, the ensemble model also captures 2.3 times as much of the mortality in its top risk decile (efficiency gain for recall10) and 1.6 times the mortality in its top risk quintile (efficiency gain for recall20). No single country always shows the highest recall10 across models, nor the lowest. Moreover, within given countries, the ensemble model outperforms the constituent models in many cases, does as well as the top one or two models in many others, and is never the worst performing model in that country. Thus, relying on the ensemble model not only yields better results than any other model on average, but also safeguards against cases where individual models

perform poorly in a given country.

Variable importance. The elastic-net logistic regression model for each country automatically chooses a subset of variables to retain in the model. Surprisingly, every variable was included in at least 23% of the country-level models. A few variables, though, are included especially often: most models included the previous death of a sibling in the first year (86%) or the death of any previous sibling at any point (76%). 77% of country level models included the child's gender; 73% included an urban/rural indicator. Over two-thirds (68%) of models included the mother's years of education, and many included the birth month (73% use sine-transformed birth month; 55% use cosine-transformed birth month). Maternal age was included in 64% of models.

For random forest, the "scaled variable importance" describes what fraction of the time a given variable was chosen to be included in the classification trees over which each random forest model aggregates. Compared to elastic-net logistic regression, random forest was somewhat more selective, with over half of variables appearing in fewer than 15% of the regression trees. A few variables were rarely included, such as clean cooking fuel (3%) and minority religion (5%). By contrast, maternal age was included nearly every time (99.6%), followed closely by wealth percentile and it's square (93%, 94%). The latter set is particularly notable, as wealth percentiles were among the *least* commonly employed in the elastic-net logit models. Other important variables included malaria incidence rate in the area (92%), age of household head (77%), and mother's age at first marriage or union (63%).

Table 4: Variable Importance

variable	**elastic survival**	**proportion positive**	**rf variable importance**
sibling death (¡1 yr.)	0.86	1.00	0.11
mother education	0.77	0.06	0.39
sin(month)	0.73	0.69	0.46
female	0.73	0.00	0.17
sibling death	0.71	1.00	0.11
rural	0.64	0.71	0.10
malaria	0.62	0.75	0.92
maternal age	0.59	0.08	1.00
prev. pregnancy ended	0.57	0.75	0.12
clean water	0.55	0.58	0.14
cos(month)	0.55	0.33	0.48
log(birth order)	0.55	0.25	0.54
mother age first union	0.52	0.45	0.63
log(num bednets)	0.52	0.09	0.38
safe floor	0.50	0.18	0.11
age household head	0.50	0.64	0.77
female household	0.50	0.55	0.11
minority religion	0.50	0.70	0.06
christian	0.35	0.43	0.08
muslim	0.35	0.71	0.06
unimproved toilet	0.32	0.86	0.09
clean cooking fuel	0.31	0.20	0.05
(wealth percentile)^2	0.27	0.00	0.94
pit toilet	0.23	0.60	0.10
wealth percentile	0.23	0.60	0.93

Discussion

Across 22 countries in sub-Saharan Africa, predicting early mortality based on wealth measures alone was ineffective: those in highest 10% and 20% risk groups accounted for only 11% and 23% of mortality respectively. Fortunately, when wealth information is combined with other risk factors, flexible models can more accurately assess individual-level risk of early mortality. On average across countries, the 10% of births with the highest risk as predicted by the ensemble model account for 22% of deaths, and the 20% at highest risk account for 35%. Compared to the wealth-only models—which we expect are optimistic benchmarks for the poverty-based targeting approach—each country's ensemble model on average would have identified 2.3 times as many deaths in the top risk decile, and 1.6 times as many in the top 20% risk group. Senegal offers a relatively representative example, in which targeting the poorest 10% will identify only 10% of all infant deaths, while targeting the 10% at highest risk according to the ensemble model would identify 29% of all deaths.

The various models used to incorporate these risk factors—a regularized logistic regression (elastic-logit), random forest (rf), extreme gradient boosting (xgb), and kernel regularized least squares (krls)—have similar average performance. However, some models exceed others in particular countries. Using the average of predictions from these (ensemble) generates better performance on average while importantly protecting against cases where individual models performed poorly in particular countries. Particularly important to many models were variables regarding the previous death of siblings in the first year, mother's age and education, malaria prevalence, and the child's gender. The value of these variables as predictors says nothing of their causal influence or importance. Rather, they are likely to be predictive not because they cause or prevent mortality, but because they signal the presence or absence of other important unmeasured factors that do.

Central to the policy-relevance of such an approach is the question of whether the variables needed to run these models are generally accessible. The risk factors we included were all chosen to be feasibly collected and employed in targeting strategies by health agencies. We avoided post-birth variables, such as birth-weight, which require real-time data. Only the sex of the child and the calendar month of birth are unknown until the birth occurs, and this would not impede the planning or siting of interventions and materials that are expected to reduce preventable deaths.

These findings are consistent with previous studies that show that births at high risk of death exist across all socioeconomic groups, and that for various purposes combined multiple risk factors and show that doing so improves our ability to identify the higher risk births.[21,11,2] Our findings are related to a large body of literature in medicine and public health that develops risk scores to identify those at risk of some event (e.g.[22,23]). To our knowledge, flexible machine learning approaches using numerous pre-birth predictors have not previously been used to develop birth-level mortality risk scores, in sub-Saharan Africa or elsewhere.

Another way to characterize these results is to construct a hypothetical scenario involving an intervention with some fixed efficacy rate and fixed budget, then examine how targeting with this system would compare to another in terms of the hypothetical lives saved. For example, consider an intervention that a given country can afford to administer to only 10% of births. Ideally, this would be targeted to the 10% at highest risk of early mortality. This intervention would not reduce mortality among children who would not have died anyway, but let us suppose it has a fixed efficacy rate, which as defined above, is the probability of saving a life that would have otherwise have been lost. In a country with N_{births} births per year, the number of lives saved per year would be efficacy $* \, Pr(\text{death}|\text{high risk}) * N_{births}/10$. Applying Bayes' rule, this is equal to efficacy $*$ recall10 $* \, Pr(\text{death}) * N_{births}$ where $Pr(\text{death})$ is the overall mortality rate expected for children born in that year (absent the intervention). Notice that the number of lives saved is proportional to the recall10 of the model used, which is why we chose this measure as a simple and policy relevant metric of performance, as opposed to other measures used for classification accuracy in machine learning such various "f-scores" or or the "area under the curve" on a receiver operator characteristic curve (AUC-ROC).

Using recent estimates of the number of births and baseline mortality rates in each of these countries,[24] the number of lives saved by an intervention with an efficacy of 5% would be 841 thousand for the wealth model, but 1.61 million for the ensemble model. In simpler terms, the efficiency gain of roughly two for these models compared to the wealth model translates into roughly twice as many lives saved by an intervention with any particular efficacy.[4] Again, we consider this a charactierization of the model performance through a hypothetical illustration, not an extrapolation intended to predict lives saved, since assumptions such as constant cost regardless of the targeting strategy or constant efficacy rate could reasonably be incorrect.

We emphasize that all models here were trained and tuned by cross-validation on a training set, and the performance indicated here represents their performance on a test set that was completely unseen during training. If such models are to be used in practice, they would be retrained on the available data in a given country at present with careful tuning by cross-validation to avoid over-fitting. Such over-fitting may be tempting for analysts but in reality leads to over-confidence in the model's ability and possibly poor performance once deployed.

Limitations and Opportunities

Our work is not without limitations. First, while our aim was to rely only on the types of predictors that are already widely available through existing IPUMS DHS data, unfortunately a few countries did not have surveys even in recent decades (Sudan) or lacked important basic variable such as district-level or lower geographic information and

[4]Note that the average efficiency gain (2.3) is not exactly the ratio of lives saved because the latter takes the relative birth rates and mortality in each country into account.

malaria prevalence (Burundi, Chad, Ethiopia, Liberia, Namibia). This limits generalization of our findings to areas with the lowest data availability, but also points to the need to identify or build suitable data sources for these countries.

Second, the DHS sampling procedure attempts to produce a sample that is representative of the population. If these are in fact not as representative as hoped, then while the results still reflect the predictive power of these models in some population of each country, the results may not be representative of these countries as a whole.

Finally, and most broadly, the central limitation is simply that not all deaths can be predicted, at least with the data we considered to be readily available in most countries. In general we find our models for the highest 10% risk group capture only 20-30% of mortality. This is beneficial and worth using for targeting since the effectiveness of an intervention is proportional to this recall rate, as shown above. However, it leaves a great deal of unpredicted mortality spread across the lower risk groups. Thus, when possible, programs must still be targeted to a much wider group in order to capture a large fraction of births that will die. Relatedly, in the future and with appropriate data, approaches like ours could be refined if information of the cause of death were available. Associating each birth to a particular cause of the death could generate models that suggest not only which births to target, but also which intervention(s) to make available to them.

References

1 Ramos AP, Weiss RE. Measuring Within and Between Group Inequality in Early-Life Mortality Over Time: A Bayesian Approach with Application to India; 2018. Https://arxiv.org/abs/1804.08570.

2 Houweling TAJ, Klaveren Dv, Das S, Azad K, Tripathy P, Manandhar D, et al. A prediction model for neonatal mortality in low- and middle-income countries: an analysis of data from population surveillance sites in India, Nepal and Bangladesh. International Journal of Epidemiology. 2019;48(1):86–198.

3 Shukla VV, Eggleston B, Ambalavanan N, et al. Predictive Modeling for Perinatal Mortality in Resource-Limited Settings. Journal of American Medical Association;3(11):e2026750.

4 Rajaratnam JK, Marcus JR, Flaxman AD, Wang H, Levin-Rector A, Dwyer L, et al. Neonatal, postneonatal, childhood, and under-5 mortality for 187 countries, 1970-2010: a systematic analysis of progress towards Millennium Development Goal 4. The Lancet. 2010.

5 Stuckler D, Basu S, Mckee M. Drivers of inequality in Millennium Development Goal progress: a statistical analysis. PLoS Medicine. 2010 Mar;7(3):e1000241.

6 Deaton A. The Great Escape: Health, Wealth, and the Origins of Inequality. Princeton University Press; 2020.

7 Bendavid E. Changes in Child Mortality Over Time Across the Wealth Gradient in Less Developing Counntries. Pediatrics. 2014;134:e1551-9.

8 Moser K, Leon D, Gwatkin D. How does Progress toward the Millennium Development Goal affects Inequality between the Poorest and the least Poor? An Analysis of Demographic and Health Survey. British Medical Journal. 2005;7526:1180-2.

9 Black R, Morris S, Bryce J, Venis S. Where and why are 10 million children dying every year? Commentary. The Lancet. 2003 February;361:2226-34.

10 Jones G, Steketee R, Black R, Bhutta Z, Morris S. How many child deaths can we prevent this year? The Lancet. 2003 July;362:65-71.

11 Ramos AP, Weiss RE, Flores M. Leave No Child Behind: Using Data from 1.7 Million Children from 67 Developing Countries to Measure Inequality Within and Between Groups of Births and to Identify Left Behind Populations. PlosOne. 2020;(7):25-45.

12 Akresh R, de Walque D, Kazianga H. Alternative Cash Transfer Delivery Mechanism: Impacts on

Routine Preventive Health Clinic Visits in Burkina Faso. National Bureau of Economic Research; 2015. 17785.
13 Akachi Y, Steenland M, Fink G. Associations between key intervention coverage and child mortality: an analysis of 241 sub-national regions of sub-Saharan Africa. International journal of epidemiology. 2018;47(3):740-51.
14 Aheto JMK. Predictive model and determinants of under-five child mortality: evidence from the 2014 Ghana demographic and health survey. BMC public health. 2019;19(1):1-10.
15 Breiman L. Random Forests. Machine Learning. 2001 October;45:5-32.
16 Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '16. New York, NY, USA: ACM; 2016. p. 785-94. Available from: http://doi.acm.org/10.1145/2939672.2939785.
17 Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and Regression Trees. Monterey, CA: Wadsworth and Brooks; 1984.
18 Hainmueller J, Hazlett C. Kernel regularized least squares: Reducing misspecification bias with a flexible and interpretable machine learning approach. Political Analysis. 2014:143-68.
19 Zou H, Hastie T. Regularization and variable selection via the Elastic Net. Journal of the Royal Statistical Society, Series B. 2005;67:301-20.
20 Krishna A. Escaping poverty and becoming poor: who gains, who loses, and why? World development. 2004;32(1):121-36.
21 Ramos AP, Heymann J, Weiss R. Improving Program Targeting to Combat Early-life Mortality by Identifying High Risk Births in India. Population Health Metrics. 2018 August;16(17).
22 Mpimbaza A, Sears D, Sserwanga A, Kigozi R, Rubahika D, Nadler A, et al. Admission Risk Score to Predict Inpatient Pediatric Mortality at Four Public Hospitals in Uganda. PLoS ONE. 2015;10(7):e0133950.
23 Beymer M, Weiss RE, Sugar C, Bourque LB, Gee GC, Morisky D, et al. Are Centers for Disease Control and Prevention Guidelines for Preexposure Prophylaxis Specific Enough? Formulation of a Personalized HIV Risk Score for Pre-Exposure Prophylaxis Initiation. Sexually Transmitted Diseases. 2017;44(1).
24 World Bank Development Indicators. Birth rates, early mortality; 2021. Data retrieved from World Development Indicators, http://data.worldbank.org/.

Beyond poverty as a proxy: reducing inequality in infant mortality by identifying and targeting higher risk births

Chad Hazlett, Antonio Ramos, and Stephen Smith

May 1, 2022

Table 1: Recall10 results, cross-validation on training set

Survey	**wealth**	**elastic-logit**	**rf**	**xgb**	**krls**	**ensemble**	**efficiency gain**
Angola 2015	0.11	0.21	0.21	0.22	0.24	0.24	2.30
Benin 2011	0.12	0.28	0.26	0.28	0.29	0.29	2.30
Burkina Faso 2010	0.09	0.16	0.15	0.16	0.14	0.18	2.08
Cameroon 2011	0.12	0.15	0.22	0.14	0.15	0.24	2.04
Congo Democratic Republic 2013-14	0.09	0.14	0.19	0.16	0.17	0.20	2.27
Cote d'Ivoire 2011	0.11	0.20	0.22	0.19	0.21	0.21	1.94
Ghana 2014	0.06	0.18	0.13	0.17	0.16	0.18	3.11
Guinea 2012	0.17	0.16	0.15	0.14	0.16	0.17	1.00
Kenya 2014	0.08	0.15	0.17	0.13	0.15	0.16	2.14
Lesotho 2014	0.07	0.09	0.16	0.08	0.15	0.14	2.00
Madagascar 2008	0.10	0.24	0.24	0.24	0.25	0.27	2.66
Malawi 2016	0.11	0.16	0.20	0.18	0.20	0.21	1.94
Mali 2012	0.12	0.20	0.22	0.21	0.21	0.24	2.11
Mozambique 2011	0.10	0.16	0.24	0.18	0.21	0.24	2.34
Niger 2012	0.10	0.15	0.20	0.16	0.15	0.19	2.02
Nigeria 2013	0.12	0.17	0.17	0.19	0.19	0.20	1.60
Rwanda 2014	0.15	0.16	0.21	0.13	0.14	0.19	1.25
Senegal 2017	0.08	0.15	0.21	0.21	0.17	0.22	2.67
Tanzania 2015	0.09	0.12	0.18	0.16	0.14	0.19	2.18
Uganda 2016	0.10	0.14	0.17	0.16	0.13	0.18	1.73
Zambia 2013	0.08	0.15	0.21	0.19	0.19	0.21	2.61
Zimbabwe 2015	0.12	0.14	0.20	0.17	0.15	0.19	1.57
Average	0.10	0.17	0.20	0.18	0.18	0.21	2.08

Table 2: Recall20 results, cross-validation on training set

Survey	**wealth**	**elastic-logit**	**rf**	**xgb**	**krls**	**ensemble**	**efficiency gain**
Angola 2015	0.24	0.39	0.37	0.40	0.38	0.42	1.76
Benin 2011	0.24	0.45	0.41	0.45	0.45	0.47	1.97
Burkina Faso 2010	0.23	0.31	0.29	0.31	0.30	0.30	1.30
Cameroon 2011	0.25	0.27	0.33	0.27	0.27	0.37	1.45
Congo Democratic Republic 2013-14	0.18	0.26	0.30	0.29	0.29	0.32	1.76
Cote d'Ivoire 2011	0.20	0.32	0.35	0.32	0.33	0.39	1.99
Ghana 2014	0.21	0.30	0.24	0.28	0.29	0.27	1.28
Guinea 2012	0.28	0.27	0.29	0.30	0.27	0.32	1.12
Kenya 2014	0.17	0.27	0.28	0.27	0.25	0.30	1.80
Lesotho 2014	0.14	0.22	0.32	0.18	0.25	0.27	1.94
Madagascar 2008	0.22	0.38	0.36	0.39	0.39	0.42	1.90
Malawi 2016	0.22	0.30	0.31	0.32	0.32	0.35	1.63
Mali 2012	0.22	0.36	0.35	0.36	0.38	0.37	1.72
Mozambique 2011	0.24	0.29	0.37	0.33	0.32	0.37	1.53
Niger 2012	0.20	0.31	0.31	0.29	0.28	0.33	1.71
Nigeria 2013	0.25	0.34	0.29	0.34	0.33	0.34	1.35
Rwanda 2014	0.26	0.36	0.30	0.27	0.31	0.32	1.23
Senegal 2017	0.20	0.30	0.31	0.33	0.31	0.34	1.68
Tanzania 2015	0.16	0.23	0.31	0.29	0.25	0.26	1.63
Uganda 2016	0.22	0.24	0.28	0.29	0.24	0.31	1.37
Zambia 2013	0.16	0.31	0.30	0.32	0.33	0.35	2.17
Zimbabwe 2015	0.23	0.31	0.34	0.32	0.29	0.33	1.47
Average	0.22	0.31	0.32	0.31	0.31	0.34	1.62

Table 3: Variable information

!

Variable	Type	Range of Values	Description
kidbordlog	Numeric	[0.6931,2.9444]	Natural logarithm of the child's birth order
pregtermin	Categorical	{NoPregTerm, PregTerm}	Ever had pregnancy terminate
agefrstmar	Numeric	[1,49]	Women's age at start of first marriage or union
hheadagehh	Numeric	[8,95]	Age of household head
urban	Categorical	{Urban, Rural}	Urban or rural living status
kidsex	Categorical	{girl, boy}	Sex of child
hheadsex	Categorical	{female, male}	Sex of household head
edyrtotal	Numeric	[1,25]	Total years educated
maternal_age_month	Numeric	[135,584]	Age of mother calculated in months
drinkwtr_new	Categorical	{bad, good}	Main source of drinking water
floor_new	Categorical	{unsafe, safe}	Main material of floor
cookfuel_new	Categorical	{unclean, clean}	Type of fuel used for cooking
toilettype_new	Categorical	{flush, pit, unimproved}	Type of toilet
religion_new	Categorical	{Buddhist, Christian, Hindu, Other}	Religion of mother
wealths	Numeric	[-16.3862,32.4577]	Wealth index categorical score
kidbirthmo	Categorical	{1,2,âŠ,12}	Month of the child's birth
kidbirthmo_sin	Numeric	[-1,1]	Sine transformation of child's birth month
kidbirthmo_cos	Numeric	[-1,1]	Cosine transformation of child's birth month
bednetnum	Numeric	{0,1,2,3,4,5,6,7+}	Number of mosquito nets owned by the household
bednetnum_log	Numeric	[0.69,2.2]	Natural logarithm of bednetnum+1
prev_death	Categorical	{no, yes}	Mother experienced previous death of child ¡12mo
prev_death_full	Categorical	{no, yes}	Mother experienced previous death of child ¡5yr
district	Categorical	Depends on the country	District where the household lived at the time of survey
Province/State	Categorical	Depends on the country	State or Province where the household lived at the time of survey
wealthp2	Categorical	{1,2,3,4,5}	Wealth quintile based on wealths (by country)
malaria_new	Numeric	[0.01,0.74]	Malaria incidence rate from year prior to child's birth
mortality.under12m	Categorical	{no, yes}	Outcome variable which records death of the child under 12mo

1 Country-by-Country Results

Efficiency gain is measured by the ratio of recall10 Results for each algorithm compared to the recall10 Results for using wealth alone.

Table 4: Detailed Results by Country

Algorithm	recall10	recall20	ROC	MRD	MRR	F1	Efficiency Gain
Angola							
Elastic Net	0.212	0.394	0.664	0.012	1.276	0.933	2.001
Ensemble	0.244	0.424	0.683	0.014	1.399	0.934	2.300
KRLS	0.236	0.384	0.669	0.016	1.379	0.934	2.227
Random Forest	0.214	0.368	0.641	0.013	1.904	0.933	2.024
Wealth	0.106	0.241	0.560	0.207	15.020	0.928	1.000
XGB	0.225	0.397	0.666	0.016	1.379	0.934	2.124
Benin							
Elastic Net	0.284	0.454	0.678	0.026	1.568	0.936	2.276
Ensemble	0.287	0.475	0.685	0.023	1.590	0.935	2.299
KRLS	0.292	0.449	0.674	0.021	1.475	0.936	2.338
Random Forest	0.260	0.411	0.662	0.022	2.046	0.935	2.083
Wealth	0.125	0.242	0.537	0.118	2.244	0.928	1.000
XGB	0.278	0.446	0.666	0.024	1.518	0.935	2.233
Burkina Faso							
Elastic Net	0.161	0.307	0.608	0.008	1.120	0.920	1.866
Ensemble	0.179	0.304	0.626	0.010	1.175	0.921	2.076
KRLS	0.145	0.297	0.611	0.010	1.152	0.919	1.673
Random Forest	0.155	0.287	0.581	0.009	1.349	0.920	1.789
Wealth	0.086	0.234	0.553	0.147	1.958	0.914	1.000
XGB	0.159	0.306	0.620	0.013	1.188	0.920	1.845
Cameroon							
Elastic Net	0.153	0.272	0.584	0.006	1.085	0.921	1.304
Ensemble	0.239	0.365	0.620	0.021	1.313	0.926	2.040
KRLS	0.146	0.266	0.584	0.005	1.074	0.920	1.247
Random Forest	0.217	0.327	0.598	0.062	1.827	0.925	1.850
Wealth	0.117	0.252	0.547	0.137	2.072	0.918	1.000
XGB	0.142	0.274	0.595	0.013	1.199	0.920	1.210
Congo Democratic Republic							
Elastic Net	0.145	0.263	0.557	0.004	1.064	0.922	1.625
Ensemble	0.202	0.325	0.604	0.008	1.162	0.926	2.266
KRLS	0.174	0.293	0.588	0.008	1.126	0.924	1.945
Random Forest	0.190	0.296	0.581	0.015	1.618	0.925	2.124
Wealth	0.089	0.185	0.520	0.064	1.503	0.919	1.000
XGB	0.162	0.294	0.590	0.007	1.116	0.923	1.821
Cote d'Ivoire							

Table 4: Detailed Results by Country *(continued)*

Algorithm	recall10	recall20	ROC	MRD	MRR	F1	Efficiency Gain
Elastic Net	0.205	0.321	0.629	0.016	1.205	0.920	1.918
Ensemble	0.208	0.395	0.654	0.018	1.241	0.919	1.943
KRLS	0.208	0.330	0.638	0.016	1.205	0.920	1.947
Random Forest	0.217	0.353	0.622	0.024	1.432	0.921	2.027
Wealth	0.107	0.199	0.529	0.076	2.132	0.911	1.000
XGB	0.193	0.324	0.619	0.015	1.185	0.919	1.806
Ghana							
Elastic Net	0.185	0.304	0.519	0.002	1.048	0.932	3.127
Ensemble	0.185	0.271	0.539	0.007	1.161	0.931	3.127
KRLS	0.165	0.290	0.567	0.000	1.009	0.931	2.796
Random Forest	0.132	0.243	0.556	0.019	1.530	0.929	2.232
Wealth	0.059	0.210	0.568	-0.038	0.889	0.926	1.000
XGB	0.172	0.285	0.539	0.006	1.135	0.931	2.901
Guinea							
Elastic Net	0.157	0.274	0.576	0.005	1.070	0.918	0.926
Ensemble	0.170	0.318	0.606	0.009	1.134	0.918	1.001
KRLS	0.157	0.274	0.582	0.005	1.071	0.918	0.925
Random Forest	0.145	0.287	0.582	0.012	1.301	0.917	0.853
Wealth	0.170	0.283	0.574	0.242	2.981	0.919	1.000
XGB	0.139	0.296	0.599	0.013	1.171	0.916	0.819
Kenya							
Elastic Net	0.148	0.271	0.524	0.001	1.025	0.932	1.948
Ensemble	0.163	0.301	0.573	0.003	1.088	0.932	2.143
KRLS	0.150	0.250	0.536	0.000	1.005	0.932	1.975
Random Forest	0.168	0.282	0.564	0.009	1.408	0.933	2.199
Wealth	0.076	0.167	0.542	-0.028	0.882	0.929	1.000
XGB	0.127	0.273	0.558	0.002	1.046	0.931	1.668
Lesotho							
Elastic Net	0.094	0.214	0.575	0.001	1.011	0.916	1.378
Ensemble	0.138	0.267	0.586	0.005	1.076	0.918	2.022
KRLS	0.145	0.249	0.548	0.001	1.017	0.920	2.133
Random Forest	0.162	0.317	0.582	0.013	1.344	0.921	2.378
Wealth	0.068	0.137	0.583	-0.026	0.859	0.914	1.000
XGB	0.077	0.181	0.522	0.003	1.047	0.915	1.133
Madagascar							
Elastic Net	0.242	0.378	0.629	0.013	1.270	0.933	2.339
Ensemble	0.274	0.424	0.656	0.024	1.473	0.934	2.652
KRLS	0.250	0.391	0.642	0.014	1.288	0.933	2.417
Random Forest	0.236	0.361	0.635	0.052	1.933	0.932	2.285
Wealth	0.103	0.223	0.549	0.126	1.532	0.925	1.000
XGB	0.245	0.391	0.653	0.016	1.338	0.933	2.364

Table 4: Detailed Results by Country *(continued)*

Algorithm	recall10	recall20	ROC	MRD	MRR	F1	Efficiency Gain
Malawi							
Elastic Net	0.164	0.301	0.582	0.004	1.086	0.931	1.508
Ensemble	0.211	0.350	0.619	0.009	1.238	0.933	1.937
KRLS	0.202	0.321	0.594	0.006	1.153	0.933	1.857
Random Forest	0.204	0.310	0.600	0.017	1.910	0.933	1.878
Wealth	0.109	0.215	0.533	0.063	1.456	0.929	1.000
XGB	0.184	0.319	0.594	0.007	1.175	0.932	1.692
Mali							
Elastic Net	0.203	0.355	0.646	0.015	1.262	0.927	1.750
Ensemble	0.245	0.371	0.666	0.025	1.410	0.929	2.114
KRLS	0.213	0.379	0.657	0.018	1.299	0.927	1.841
Random Forest	0.221	0.347	0.634	0.050	1.742	0.928	1.909
Wealth	0.116	0.216	0.555	0.181	3.474	0.921	1.000
XGB	0.208	0.358	0.654	0.017	1.291	0.927	1.795
Mozambique							
Elastic Net	0.156	0.286	0.567	0.005	1.077	0.921	1.523
Ensemble	0.240	0.370	0.631	0.020	1.296	0.926	2.341
KRLS	0.207	0.323	0.591	0.010	1.148	0.924	2.023
Random Forest	0.244	0.365	0.625	0.051	1.676	0.927	2.386
Wealth	0.102	0.242	0.536	0.089	2.105	0.917	1.000
XGB	0.179	0.326	0.598	0.015	1.228	0.922	1.750
Niger							
Elastic Net	0.149	0.307	0.612	0.008	1.127	0.923	1.554
Ensemble	0.194	0.335	0.635	0.018	1.287	0.925	2.021
KRLS	0.149	0.282	0.625	0.009	1.154	0.923	1.554
Random Forest	0.205	0.307	0.610	0.046	1.671	0.927	2.131
Wealth	0.096	0.196	0.532	0.158	2.954	0.919	1.000
XGB	0.160	0.286	0.607	0.009	1.147	0.924	1.665
Nigeria							
Elastic Net	0.173	0.338	0.621	0.012	1.171	0.920	1.401
Ensemble	0.199	0.339	0.634	0.017	1.227	0.922	1.604
KRLS	0.186	0.332	0.623	0.012	1.164	0.921	1.502
Random Forest	0.175	0.294	0.603	0.031	1.375	0.920	1.413
Wealth	0.124	0.251	0.568	0.223	1.934	0.916	1.000
XGB	0.189	0.342	0.622	0.012	1.175	0.921	1.526
Rwanda							
Elastic Net	0.164	0.357	0.623	0.004	1.114	0.936	1.103
Ensemble	0.186	0.320	0.618	0.005	1.172	0.936	1.251
KRLS	0.142	0.312	0.609	0.003	1.099	0.935	0.952
Random Forest	0.207	0.296	0.578	0.010	1.517	0.938	1.387
Wealth	0.149	0.259	0.561	0.196	5.841	0.936	1.000

Table 4: Detailed Results by Country *(continued)*

Algorithm	recall10	recall20	ROC	MRD	MRR	F1	Efficiency Gain
XGB	0.126	0.266	0.556	0.003	1.090	0.935	0.845
Senegal							
Elastic Net	0.149	0.299	0.580	0.004	1.101	0.930	1.773
Ensemble	0.224	0.339	0.609	0.033	1.810	0.933	2.660
KRLS	0.165	0.315	0.581	0.004	1.100	0.931	1.957
Random Forest	0.209	0.311	0.587	0.059	2.194	0.933	2.475
Wealth	0.084	0.203	0.535	0.106	1.687	0.927	1.000
XGB	0.209	0.333	0.595	0.064	3.297	0.933	2.476
Tanzania							
Elastic Net	0.121	0.226	0.561	0.002	1.054	0.930	1.397
Ensemble	0.188	0.262	0.582	0.005	1.145	0.932	2.169
KRLS	0.144	0.254	0.549	0.004	1.098	0.931	1.667
Random Forest	0.180	0.309	0.584	0.011	1.459	0.932	2.080
Wealth	0.086	0.161	0.545	-0.128	0.251	0.928	1.000
XGB	0.164	0.285	0.560	0.005	1.110	0.932	1.895
Uganda							
Elastic Net	0.145	0.240	0.536	0.001	1.019	0.930	1.402
Ensemble	0.178	0.307	0.579	0.002	1.059	0.932	1.725
KRLS	0.134	0.237	0.535	0.001	1.027	0.930	1.299
Random Forest	0.167	0.278	0.554	0.001	1.620	0.931	1.624
Wealth	0.103	0.224	0.537	-0.021	0.919	0.929	1.000
XGB	0.157	0.294	0.576	0.004	1.104	0.931	1.526
Zambia							
Elastic Net	0.150	0.310	0.579	0.005	1.099	0.928	1.875
Ensemble	0.209	0.354	0.617	0.008	1.190	0.930	2.619
KRLS	0.186	0.333	0.588	0.005	1.098	0.930	2.327
Random Forest	0.209	0.300	0.598	0.016	1.588	0.931	2.618
Wealth	0.080	0.163	0.529	-0.009	0.990	0.924	1.000
XGB	0.194	0.320	0.595	0.008	1.160	0.930	2.422
Zimbabwe							
Elastic Net	0.143	0.308	0.576	0.005	1.095	0.926	1.176
Ensemble	0.190	0.333	0.642	0.010	1.216	0.929	1.569
KRLS	0.154	0.292	0.600	0.005	1.108	0.927	1.265
Random Forest	0.201	0.338	0.629	0.019	1.726	0.930	1.653
Wealth	0.121	0.228	0.570	2.029	-170.854	0.925	1.000
XGB	0.175	0.318	0.588	0.009	1.184	0.928	1.441