BMJ Open

Testing the socioeconomic and environmental determinants of better child-health outcomes in Africa: a cross-sectional study among nations

Corey J A Bradshaw [ID],[1] Sarah P Otto,[2] Alicia A Annamalay,[3] Sam Heft-Neal,[4] Zachary Wagner,[5] Peter N Le Souëf[3]

ABSTRACT

Objective We sought to test hypotheses regarding the principal correlates of child-health performance among African nations based on previous evidence collected at finer spatial scales.

Design Retrospective, cross-sectional study.

Setting All countries in Africa, excluding small-island nations.

Primary and secondary outcome measures We defined a composite child-health indicator for each country comprising the incidence of stunting, deaths from respiratory disease, deaths from diarrhoeal disease, deaths from other infectious disease and deaths from injuries for children aged under 5 years. We also compiled national-level data for Africa to test the effects of country-level water quality, air pollution, food supply, breast feeding, environmental performance, per capita wealth, healthcare investment, population density and governance quality on the child-health indicator.

Results Across nations, child health was lowest when water quality, improved sanitation, air quality and environmental performance were lowest. There was also an important decline in child health as household size (a proxy for population density) increased. The remaining variables had only weak effects, but in the directions we hypothesised.

Conclusions These results emphasise the importance of continued investment in clean water and sanitation services, measures to improve air quality and efforts to restrict further environmental degradation, to promote the UN's Sustainable Development Goal 3 target to '… end preventable deaths of newborns and children under 5' and Goal 6 to '… ensure access to water and sanitation for all' by 2030.

INTRODUCTION

WHO estimates that 5.6 million children under the age of 5 died in 2016, with over half of these deaths deemed preventable and/or treatable with minimal intervention.[1] While child survival has improved substantially since 1990 (56% decline in mortality), children under 5 in sub-Saharan Africa are still 15 times more likely to die than those in

Strengths and limitations of this study

► The first continental-wide analysis of overall child-health performance among African nations.
► Poorer child health occurs in countries with the lowest availability of clean water and sanitation, highest household size, greatest incidence of poverty and highest environmental degradation.
► This represents the first empirical demonstration that broad-scale environmental degradation and high household size deteriorate child health in low- and middle-income nations.
► A limitation is that our country-scale assessment cannot identify subnational patterns of success of project-based health interventions.
► These relationships are derived from the most recent data and therefore do not necessarily reflect past child-health outcomes and socioeconomic conditions that might have been historically more important than those today.

high-income countries, with 1 in 13 children there dying before their fifth birthday.[1] Still, many African nations have made recent progress in improving child health,[2 3] with under-5 mortality declining from 176 to 90 per 1000 live births (a 49% reduction) since 1990 (to 2013).[4]

Nonetheless, increasing pressure from climate change and environmental deterioration is potentially weakening interventions to improve child health in this region.[5] For example, most of the global burden of disability and mortality related to climate change occurs in children under 5.[6–8] Africa also has the world's highest rates of projected population growth,[9 10] with little indication that fertility rates there are decreasing as rapidly as they have in other regions of low- and middle-income countries over the last few decades.[11] The observation that regions with high population growth rates have more poverty and poorer child health, combined

with increasing environmental degradation and biodiversity loss globally,[12] suggests that young children in many African nations stand to suffer the worst consequences of these trends. Indeed, as of 2015 and the end of the period covered by the Millennium Development Goals, only eight African countries were on track to achieve Target 4A (reducing under-5 mortality by two-thirds between 1990 and 2015).[4] If more African nations have any chance of meeting the post-2015 Sustainable Development Goal 3 target to '… end preventable deaths of newborns and children under 5' by 2030, we need to evaluate the conditions leading to better child-health outcomes at broad spatial scales.[8]

While there are many studies examining the conditions, interventions and policies that correlate with child-health outcomes in Africa, most are applied at subnational scales or focus on single regions or nations,[13–15] and most examine one or only a few hypothesised variables at a time[2 5 8 13 16–18] (online supplementary information, table S1). Indeed, a recent study mapping child growth in Africa between 2000 and 2015 identified subnational trends in stunting, wasting and underweight prevalence, but did not directly test any plausible correlates potentially driving the spatial and temporal trends.[2] As such, it is unclear if generalisable 'rules of thumb' regarding the relative effectiveness of different interventions can be formulated across greater spatial scales. Furthermore, even the concept of 'child health' depends on the metrics one chooses to employ as a quantitative measure of outcomes: from raw mortality rates, to the incidence and prevalence of different infectious diseases, to measures of undernutrition such as stunting (online supplementary information, table S1). Furthermore, interventions that improve child health according to one metric might not always lead to improvements in others. Therefore, our aim was to construct models examining the relationships between current environmental and socioeconomic conditions and combined metrics of child health across the entire African region, such that future policy interventions can be prioritised according to their relative contribution to reduce future child suffering and death. To this end, we applied the most recent country-level data across several metrics and correlates to explain variation in child health among African nations.

Our aim was to test hypotheses regarding the principal correlates of child-health performance among African nations based on previous evidence collected at finer spatial scales. We first hypothesised that increasing access to improved water and sanitation improves child health based on the evidence that such access lowers the probability of death, morbidity, undernutrition and stunting in young children.[13 16 17 19 20] Next, we hypothesised that increasing food supply improves child health given that the quantity and quality of food available has been identified as one of the main drivers of reductions in child stunting.[2 21] Likewise, we hypothesised that exclusively breast feeding during the first 6 months of life improves the long-term health prospects of these children.[22–25]

Given the evidence that access to green space improves child health in high-income countries,[26–28] we used an index of environmental performance for African nations[29] to test the hypothesis that declining environmental conditions compromise child health. Given the recent strong evidence that poor air quality is responsible for substantial infant deaths in Africa,[5 18] we also hypothesised that air pollution at the national scale would lead to poorer overall child health.

We also examined several socioeconomic components potentially linked to child-health performance. We hypothesised that a country's capacity to invest in the health of its citizens is correlated with per capita wealth, as has been found among regions in China[30]; thus, we predict that wealthier African nations have better child-health outcomes on average than poorer nations. As a more direct measure of a nation's investment in health infrastructure and services, we hypothesised a positive correlation between total per capita health investment (both domestic and aid-based) and child health.[31 32] Likewise, we hypothesised that a country's governance quality would be correlated with average child health given the evidence that increasing democracy, and hence, reducing corruption, generally improves child health outcomes.[33] Finally, we hypothesised that demographic competition could in itself compromise child health given the negative relationships observed between health performance (children and adults) and household size (or number of siblings and/or population density);[14 15 19 34–36] we therefore predicted a negative correlation between child health and household size for African nations. Collectively then, our combined national-scale metrics of child health performance in Africa is the first comprehensive study of the multiple and potentially interacting correlates of poor child-health outcomes in this region.

METHODS
Child health data
Our first goal was to define a composite child-health indicator for each nation on the African continent. While there are many ways to measure a nation's overall health performance, we focused on measures with substantial regional resolution to link health outcomes better to the specific environmental histories of regions. We define 'child' here following the WHO standard of birth to 5 years of age. We reasoned that the following available indicators would provide a highly informative composite child-health index for African nations:

1. *Stunting (STUNT)*: stunting is an index of undernutrition that measures the degree to which a child's growth is limited by lack of food.[2] We accessed these country-specific data for the most current period (range: 1994–2015) from the WHO Data Repository (who.int), which defined stunting as the percentage of children under age 5 years falling below 2 standard deviations of the median height-for-age of the WHO standard.

2. *Deaths from acute lower respiratory infections*: we acquired the number of deaths per 1000 live births resulting from acute lower respiratory infections for children aged from 1 to 59 months for 2015 across all African nations from WHO.

3. *Deaths from diarrhoeal disease*: we acquired the number of deaths per 1000 live births resulting from diarrhoeal disease for children aged from 1 to 59 months for 2015 across all African nations from WHO.

4. *Deaths from infectious diseases*: we acquired the number of deaths from three different infectious diseases for children aged from 1 to 59 months for 2015 from WHO: malaria, HIV/AIDS and pertussis (because deaths from acute lower respiratory infections do not include those arising from pertussis). We combined all deaths from these sources and calculated the proportion of 'infectious disease' deaths per total live births per African nation.

5. *Deaths from injuries*: we obtained the number of deaths per 1000 live births resulting from injuries for children aged from 1 to 59 months for 2015 across all African nations from WHO.

Combined child-health index

We ordered countries for each of the five child-health variables from the lowest health (rank=1) to highest health (maximum number of countries in list) using the *rank* function (means averaged) in the R environment[37] (ie, we did not consider the magnitude of the differences among absolute values between countries). To construct a mean rank across all five variables, we calculated geometric mean rankings for countries[38] to provide a measure of relative distance between countries in the final composite rank. This ranking approach down-weights countries that score poorly in any one of the metrics, even if they do well in others:

$$\text{geometric mean rank} = 10^{\sum_{i=1}^{k} \frac{\log_{10}(\text{rank}(x_i))}{k}},$$

where x_i=health metric i (for k metrics considered). The geometric mean is also highly correlated (Spearman's ρ=0.99) with the arithmetic mean rank among countries.

Environmental and socioeconomic data

As we outlined in the 'Introduction' section, we first hypothesised that increasing access to water quality and sanitation reduces the incidence of childhood diseases and mortality.[13 16 17 19 20] From the World Bank database (data.worldbank.org), we obtained estimates of the proportion of each country's population with access to (*i*) *improved water sources* and (*ii*) *improved sanitation facilities*. We calculated the mean proportion for these two variables, and we then scaled and centred the proportions to provide a 'water health' index across countries using the *scale* function in R. The availability of food is also an important determinant of child health globally,[21] so we hypothesised that a greater relative food supply per capita would elicit better child health. Thus, we obtained the daily available caloric value per capita (kcal/person/day)

from the Food and Agriculture Organization (FAO) as an index of *food supply*, scaling and centring the variable for subsequent analyses. Similarly, breast feeding improves the health prospects for infants even well beyond the breastfeeding period itself,[22–25] especially when access to clean water, sewage disposal and adequate fuel supply is deficient. Accessing data from WHO, we summarised the *percentage of infants exclusively breast fed* for the first 6 months of life for the most recent years available (2002–2016, depending on country), scaling and centring the variable for subsequent analyses.

Poor *air quality* was responsible for a mean of 22% of infant deaths among 30 sub-Saharan countries in Africa,[5] and ambient particulate matter is associated with child and maternal mortality across all of Africa where exposure exceeds the WHO standards[18]; we therefore hypothesised that it would also explain variation in child health among countries in our sample. We use satellite-derived data on particulate matter <2.5 µm in diameter (PM2.5) concentrations compiled by the Atmospheric Composition Analysis Group at Dalhousie University, consisting of annual, bias-corrected, average surface PM2.5 concentrations at 0.1×0.1° spatial resolution with global coverage.[39] These data are derived from several satellite-based, atmospheric optical depth-measurement instruments combined with aerosol profile measurements, and satellite, weather and seasonality data from the GEOS-Chem Chemical transport model to quantify the relationship between column aerosol optical depth and surface PM2.5, measured at available ground-based stations. To derive nationally representative estimates of air quality, we overlapped the PM2.5 grid cells on national boundaries[40] and then aggregated them at the country level using a high-resolution spatial weighting of human population distribution (ie, as a better proxy for exposure than an overall country mean) from the Landscan 2015 dataset.[41]

Next, we used a country's relative *environmental performance index*[29] as a proxy for environmental degradation. We hypothesised that child health might be lower when a country's environment is relatively more compromised, using an index that combines each country's ecological footprint,[42] megafauna conservation index,[43] proportion of species threatened (iucnredlist.org), freshwater removals (data.worldbank.org), forest loss,[44] livestock per hectare of arable land (fao.org/faostat), croplands extent (data.worldbank.org) and CO_2-e emissions (data.worldbank.org).

As indices of a country's capacity to invest in the health of its citizens, we accessed each country's per capita *gross domestic product* (GDP) (corrected for purchasing-power parity) as an index of total wealth from the World Bank database. Some countries were missing GDP estimates for certain years, so we took the mean of values from 2011 to 2015 as an indication of mean per capita GDP to maximise the sample size of countries considered. We hypothesised that the higher a country's economic resources, the greater would be its citizens' ability to afford food, technologies and medicines that would lower child diseases

and mortality resources on average.[30] More specifically, we added (*i*) the domestic per capita health expenditure (US$) for the latest year available (2014) from the World Bank database and (*ii*) the total Development Assistance for Health (DAH) from 2005 to 2015 (expressed in 2015 US$ per capita) from the Institute for Health Metrics and Evaluation DAH Database[45] as an index of foreign health aid/investment. We scaled and centred the $\log_{10}$ of the sum of these two per capita values as an index of total *health investment* (ie, both domestic and foreign).

We next hypothesised that poorer overall governance would lead to poorer child-health outcomes.[33] For this test, we used the *overall governance score* from the 2015 Ibrahim Index of African Governance,[46] which includes measures of safety and rule of law, participation and human rights, sustainable economic opportunity and human development indicators in its normalised overall score (which we subsequently scaled and centred for all analyses). Finally, we accessed the United Nations Population Division Household and Composition (2017) database (un.org) for the most up-to-date (1994–2015, median=2013) estimate of *household size* (mean number of people per household) as an index of proximate population density on child living conditions. We hypothesised that increasing household size might lead to greater pressure on per-household resources,[47] thus lowering child health.[14 15 19 34–36] Household size is likely a better proxy for resource competition than country-level population density because the former is a more direct measure of the conditions that directly influence child health. For example, high-dependency households (ie, total number of children aged <16 years in a household) had higher proportions of stunted and underweight children than lower-density households in Ghana,[14] with the effect of the number of children in the household mediated by nucleation (ie, tendency towards households comprising two parents and their own children) and conditioned on the wealth status of the household.[14] Furthermore, household crowding increased mortality from measles in West African communities.[48] These phenomena likely arise because in extended-family households, resources are more likely to flow from children to adults, whereas in more nucleated families the resource flow tends to flow more from parents to children.[49]

Boosted regression trees

We applied a machine-learning technique—boosted-regression trees[50]—implemented using the `dismo` R library and its function *gbm.step*,[51] setting the error distribution family as Gaussian (scaling and centring all metrics ensured Gaussian error behaviour and an identity link between the predictors and the response), the bag fraction to 0.80, the learning rate to 0.0001, the tolerance to 0.0001 and the tree complexity to 2. The latter setting considers only first-order interactions and combines these effects if present into the relative influence scores (see below). Boosted regression trees are particularly suited to the data and relationships we hypothesised because such

hybrid statistical/machine-learning techniques do not assume an appropriate data model to build the regression trees like purely statistical techniques do[50]; instead, they apply an algorithm to learn the relationship between the predictors (socioeconomic variables) and the response (child health).[52] This places emphasis on predictive performance rather than on the underlying model *per se*.[50] Other advantages of boosted regression trees over many other statistical techniques is that they can easily combine predictor variables of any type, results are unaffected by different scales of measurement, results are insensitive to outliners, they can accommodate missing data in the predictors, they automatically model interactions and the fitting process is stochastic such that it generally improves predictive performance of the final model.[50 53]

To assess the relative effect of each socioeconomic/environmental variable on the composite child-health metric, and subsequently on each individual child-health variable, we calculated the boosted-regression tree metrics of relative influence.[54] Relative influence I is defined as the relative influences of the individual inputs x_j on the variation of the function $\hat{F}(x)$, the latter being the function that maps the explanatory variables x to the response variable y. For a collection of M decision (regression) trees $\{T_m\}_1^M$, the squared influence $\hat{I}^2$ is calculated as:

$$\hat{I}^2 = \frac{1}{M}\sum_{m=1}^{M}\hat{I}_j^2(T_m),$$

where $\hat{I}_j^2(T) = \sum_{t=1}^{j-1}\hat{i}_t^2(v_t = j)$, which is the summation over the non-terminal nodes t of the J-terminal node tree T, v_t is the splitting variable associated with node t and $\hat{i}_t^2$ is the corresponding empirical improvement in squared error resulting from the split.[55] These squared influences then sum to 100 over all x explanatory parameters, which is analogous to the percentage of variation in the response variable y explained by each parameter. We also plotted the predicted values from the final tree against increments of the transformed (centred and scaled) individual response variables. We assessed goodness-of-fit of each final tree structure using the percentage of deviance explained (D^2 and the cross-validation correlation coefficient β_{CV} (± its standard error calculated over all tree iterations).

To quantify uncertainty in the relationships between the predicted child-health values and each contributing predictor, we also resampled the dataset with replacement 1000 times, with each resampled dataset equating to the size (*n* countries) of the original. We passed each resampled dataset to the boosted regression tree (same parameters as the original) and then calculated the 2.5th and 97.5th percentiles of the respective distribution for each predicted value as the uncertainty bounds. We also applied a kappa (κ) limitation to the resampled selections, whereby we only retained the resampled mean ranks within $\kappa\sigma$ of the overall average mean (here we set κ=2), thereafter recalculating the average and standard deviation of the mean rank, and repeating the process five times. We used this iterative $\kappa\sigma$ 'clipping'

approach[56]—which is often used in image processing to remove artefacts when stacking subframes[57]—to limit the influence of outliers in estimating the range of mean rank across all 1000 iterations.

Additional assumption-checking models

To account for potential intercorrelations among hypothesised socioeconomic and explanatory variables,[38] we also applied structural (path) equation models to account for the hypothesised relationships[58] (online supplementary information); however, these models universally had poor goodness-of-fit, even though they essentially supported the results derived from the boosted regression trees (see online supplementary information and 'Results' section). In addition to the structural equation models, we employed general linear mixed-effects models (GLMM) with a 'region' random effect according to two different socioeconomic regionalisations (groupings) for the African continent to account for any broad-scale spatial non-independence. Although including all data in this way ignores other non-independence issues (eg, country adjacencies), it identifies which correlates are likely to provide some explanatory power. Model results were broadly similar between the GLMM and boosted regression trees, although there were some differences in terms of relative ranking of the most influential predictors (online supplementary information). However, given the lack of a strong goodness-of-fit for the structural equation models, and the clear advantages of boosted regression trees over GLMM in particular, we contend that the boosted regression trees used to rank the relative importance of the predictors considered provide more robust and credible results than those derived from the other methods. The results of the structural equation models and GLMM merely provide additional confidence that the results arising from the boosted regression trees are not compromised by potential spatial autocorrelation (GLMM) or explicit relationships between predictor variable (structural equation models).

Patient and public involvement statement

Neither patients nor the public was involved in this study.

RESULTS

Composite child-health index

A non-parametric (Kendall's τ) correlation matrix among the component child-health metrics demonstrated moderate ($0.280 \leq \tau \leq 0.850$) positive relationships among the ranked variables (table 1). This suggests that a geometric mean rank provided a good central tendency of overall child health for the boosted regression tree models (see below).

Of the 10 countries with the healthiest children, 5 were in the far north (Tunisia, Libya, Morocco, Algeria, Egypt) (table 2; figure 1). The next-highest regional concentration of relatively child-healthy countries were in west-central Africa (Gabon, Congo), West Africa (Senegal

Table 1 Correlation (Kendall's τ) matrix of child health component variable ranks

	STUNT	RESP	DIAR	INFE
RESP	0.294			
DIAR	0.306	0.850		
INFE	0.280	0.538	0.552	
INJR	0.365	0.720	0.796	0.431

INJU, injuries[44]; DIAR, diarrhoeal disease; INFE, infectious diseases; RESP, respiratory infection[43]; STUNT, stunting.

and Ghana), East Africa (Kenya) and in the far south (Botswana, Namibia, South Africa) (table 2; figure 1). The countries with the worst overall child-health performance were Angola, Chad, Central African Republic, Sierra Leone, Somalia and Mali (table 2; figure 1).

Boosted regression trees

The boosted regression trees supported all of our hypotheses, although the lower the relative influence, the higher the uncertainty in the predicted relationships (figure 2). The full dataset for the composite child-health model used a total of 40 100 trees (% deviance explained D^2=66.6%; cross-validation correlation coefficient β_{CV}=61.1%±12.0%), with the highest relative contribution to the variation in the composite child-health index observed among countries coming from access to improved water and sanitation sources, followed by per capita GDP, household size, air pollution and environmental performance (figure 2A). For these four predictors, increasing the predictor by 1 SD (as measured among the countries) resulted in a mean relative increase in child health of 20%, 15% and 9% for GDP, water quality and environmental performance, respectively (ignoring non-linearities for each variable; percentages of equivalent unit change). Similarly, increasing the household size data by 1 SD resulted in a mean decrease in child health of 6%.

The κ-clipped resampling procedure generally supported the relative contributions (figure 2A, error bars). Furthermore, as the relative influence of a particular predictor declined, the uncertainty window for the predicted child-health values widened (figure 2). Here, increasing per capita wealth (GDP) was positively correlated with child health (figure 2B), and the greater the proportion of a country's population with access to clean water and sanitation, the lower the overall burden of disease of children (figure 2C). Increasing household size was negatively correlated with child health (figure 2D), as was declining environmental performance (figure 2E). Air pollution as measured by PM2.5 indicated a negative relationship to child health (figure 2F). The remaining variables followed the hypothesised relationships, although their uncertainties were considerably higher (figure 2G–J). The greater the percentage of infants exclusively breast fed for the first 6 months of life (figure 2G) and the higher a country's governance quality

Table 2 Ranking results (*n*=49 countries)

Country	ISO	STUNT	RESP	DIAR	INFE	INJU	CHLTH$_{gm}$
Tunisia	TUN	49	48.5	48	49	49	48.70
Libya	LBY	43	48.5	49	47	46	46.65
Morocco	MAR	47	47	46	46	45	46.19
Algeria	DZA	48	45	45	45	47	45.98
Egypt	EGY	40	46	47	48	48	45.70
Gabon	GAB	46	39	43	28	44	39.41
Congo	COG	42	42	41	30	42	39.08
Botswana	BWA	26	44	44	42	43	39.06
Senegal	SEN	44	38	35	38	36	38.08
Kenya	KEN	34	40.5	40	29	41	36.58
Namibia	NAM	39	31	37	37	37	36.09
South Africa	ZAF	38	35	39	34	34	35.94
Ghana	GHA	45	40.5	36	22	40	35.67
Djibouti	DJI	21	36	29	40	38.5	32.04
Tanzania	TZA	17	37	38	33	31	30.04
Gambia	GMB	37	31	27	32	21	29.08
Uganda	UGA	19	31	33	23	32	26.98
Rwanda	RWA	4	43	42	39	38.5	25.53
Ethiopia	ETH	10	29	31	41	27	25.09
Swaziland	SWZ	35	25	25	25	16	24.46
Zimbabwe	ZWE	30.5	26	23	24	17	23.68
Liberia	LBR	24	20.5	26	18	28	23.01
Malawi	MWI	8	33	30	20	26	21.04
Eq Guinea	GNQ	33	14	21	12	33	20.74
Mauritania	MRT	41	19	12	31	13	20.66
Zambia	ZMB	11	27	28	21	20	20.35
Madagascar	MDG	3	34	32	36	29	20.25
Togo	TGO	32	20.5	24	8	23	19.61
Côte d'Ivoire	CIV	28	18	22	7	35	19.36
Guinea Bissau	GNB	30.5	12	14	26	19	19.08
Eritrea	ERI	2	28	34	44	30	19.06
Sudan	SDN	14	23	19	43	8	18.39
South Sudan	SSD	27	7	17	17	24	16.73
Lesotho	LSO	22	11	13	19	15	15.51
Mozambique	MOZ	5	24	20	9	22	13.66
Burkina Faso	BFA	16	22	18	4	18	13.55
Cameroon	CMR	25	16	11	13	7	13.20
Guinea	GIN	20	9.5	16	5	25	13.06
Bénin	BEN	18	9.5	7.5	11	9	10.49
Dep Rep Congo	COD	7	13	10	10	11	10.00
Nigeria	NGA	23	6	6	6	12	9.02
Burundi	BDI	1	17	15	27	5.5	8.23
Niger	NER	6	5	9	16	5.5	7.50
Mali	MLI	13	15	7.5	1	14	7.28
Somalia	SOM	36	1	2	35	2	5.50

Continued

Table 2 Continued

Country	ISO	STUNT	RESP	DIAR	INFE	INJU	CHLTH$_{gm}$
Sierra Leone	SLE	15	8	5	2	4	5.45
Centr Afr Rep	CAF	9	4	4	3	10	5.33
Chad	TCD	12	2	3	15	3	5.04
Angola	AGO	29	3	1	14	1	4.14

CHLTH$_{gm}$, geometric mean rank of composite child-health index (highest=healthiest); DIAR, diarrhoeal disease; INFE, infectious diseases; INJU, injuries[44]; ISO, alpha-3 country code; RESP, respiratory infection[43]; STUNT, stunting.

(figure 2H), the better the outcomes for child health. Investment in healthcare and services also demonstrated only a weak improvement in child health (figure 2I), whereas there was a non-linear response of child health to per capita food supply (per capita calories), with child health worsening initially with increasing food supply and then improving again (figure 2J).

Breaking down the composite child-health index into its component variables and re-running the boosted regression trees for each separately confirmed the overall pattern of relative contribution to explaining the variance among countries (table 3; figure 3A), although the median relative contributions of each predictive variable were more even (figure 3A). Taking the median relative contribution among all five health variables changed the order of variable importance determined by the composite child-health index for most variables (figures 2A and 3A). Most individual health indicators increased with increasing air pollution apart from stunting (figure 3B), and all variables improved with greater access to improved water and sanitation (figure 3C). Stunting decreased most strongly with increasing per capita wealth, followed by a decrease
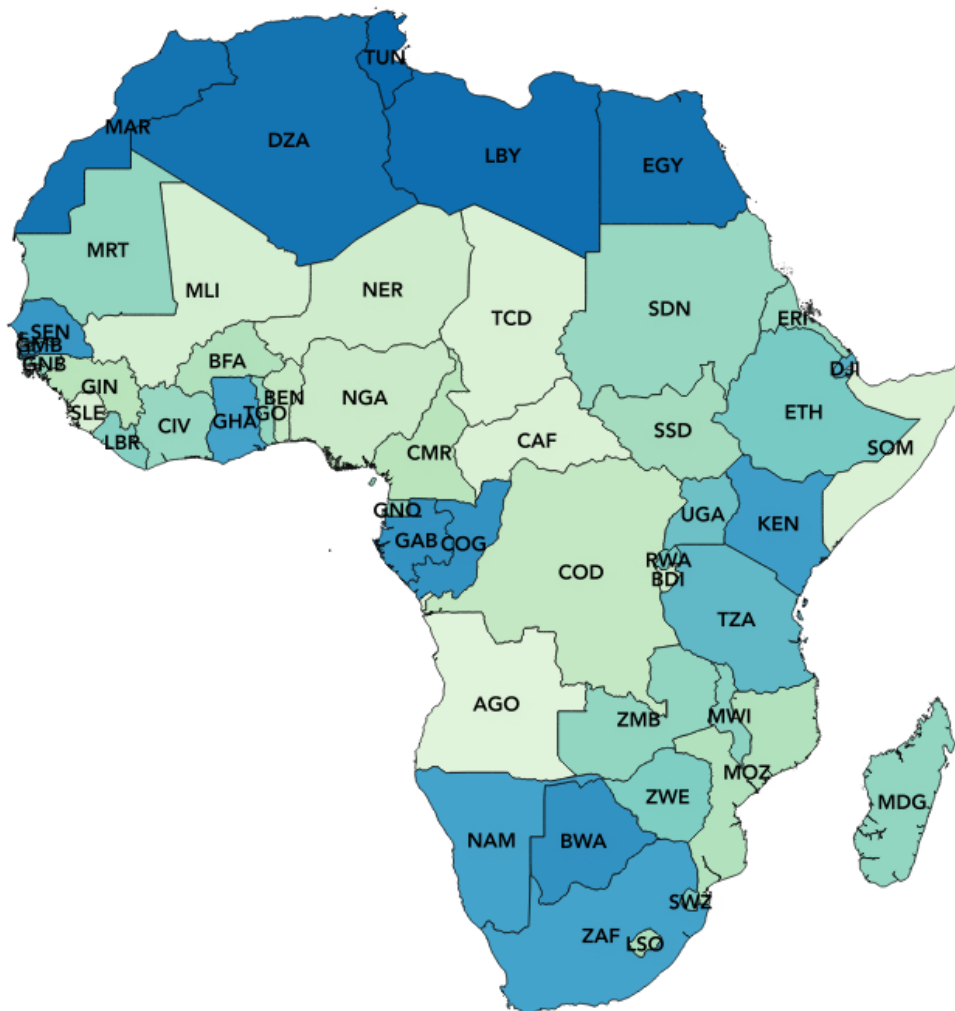

Figure 1 Map of countries in Africa shaded according to the geometric mean rank of the composite child-health index (blue=healthiest; green=unhealthiest; see table 2 for values). Geographic data from maplibrary.org (public domain).

in infectious diseases (figure 3D). Likewise, the incidence of all individual variables declined with increasing environmental performance (weakest for infectious diseases) (figure 3E). The proportion of infants exclusively breast fed also improved most child health metrics, except interestingly the incidence of stunting (and injuries) (figure 3F), whereas improving governance improved all health metrics except for infectious diseases and stunting (figure 3G). The individual relationships with food supply were highly variable (figure 3H), demonstrating why the composite health index was highly non-linear (figure 2J). Health investment had weak or no effect (figure 3I), and the effect of household size was weak for all variables except injuries and diarrhoea (figure 3J). These results are all broadly consistent between the boosted regression trees, structural equation models (online supplementary information tables S2–S7) and the general linear mixed-effects models (online supplementary information, figure S1, tables S8 and S9) (ie, the dominant predictors appeared within the main, top-ranked models for each method), although some differences in predictor rankings were apparent (online supplementary information).

DISCUSSION

There was substantial regional clustering of African nations in the overall child-health metric, with higher relative performance in the far north and south, and pockets of higher performance in western and eastern Africa (figure 1). Like recent work on spatial variation of child undernutrition[2] and mortality from air pollution,[5 18] the Sahel region (especially West Africa) generally had some of the poorest child-health outcomes relative to elsewhere in the continent, although there were poor performers in nearly all regions except the far north.

All of the relationships we hypothesised were supported based on the three modelling approaches, although the dominant correlates explaining the most variation in overall national-scale child health statistics were a country's per capita wealth (as measured by per capita GDP), access to improved water and sanitation, household size, environmental performance and air quality (figure 2). All other correlates were comparatively less important for explaining the national-scale spatial variance in child-health metrics among African nations, even though these too mostly revealed relationships in the predicted directions. We stress that these relationships are derived from the most recent data and therefore represent the present snapshot in time; thus, they do not necessarily reflect past child health outcomes and socioeconomic conditions in African nations[2] that might have been historically more important than those today (eg, food supply). Nor can our data identify potentially important within-country (subnational) relationships that might operate at finer spatial scales (online supplementary information table S1). Of course, another limitation of an analysis of this spatial scale is that even with statistical evidence for the hypothesised relationships, we cannot necessarily infer causality. However, that nearly all of the hypothesised relationships derived mainly from smaller-scale studies examining single correlates were supported at the continental scale contributes to the strength of inference of consistent responses across multiple spatial scales.

Our results nonetheless emphasise the importance of continued investment in clean water and sanitation services, measures to improve air quality and efforts to restrict further environmental degradation to maximise the rate at which African nations meet the Sustainable Development Goal 3 target to '… end preventable deaths of newborns and children under 5' and Goal 6 to '… ensure access to water and sanitation for all' by 2030. Indeed, even after World War I in Europe, improvements in sanitation, access to clean water (and nutrition and child care) were the principal drivers of improved child survival,[17] and the evidence from many low- and middle-income nations,[19] and African countries in particular,[13 16] is that improving access to clean water, sanitation and high air quality are among some of the most effective tools to improve child health. These actions are all the more imperative considering that even exposure to 'low' concentrations of many pollutants and pathogens in water is now regarded as dangerous.[20]

A more general development goal to improve sustainable economies and societies (eg, Sustainable Development Goals 1: no poverty, 2: zero hunger, 8: decent work and economic growth, 12: responsible consumption and growth) measured by raising the per capita wealth of citizens of African nations (crudely measured by GDP), was also a strong predictor of child-health performance. Indeed, the wealthiest African countries as measured by per capita GDP had the best child-health outcomes on average. However, the relationship between wealth and health is at least partially circular, given that the status of health during childhood is an important determinant of productivity and earning prospects as an adult[34] and that at the scale of entire societies, improved health has a positive effect on economic growth.[59] Thus, failing to break out of the poverty trap is partially a result of poor health begetting lower economic performance, which itself erodes health outcomes.[60] In fact, it has been estimated that the African region will lose approximately 6% of its 'non-health' GDP (ie, the product of the number of discounted life years above the minimum employment age lost, per capita non-health GDP in purchasing power parity and total child deaths) from the future years of life lost from 3 million child deaths in 2013.[4]

Dense living conditions can contribute to poor health outcomes for children, especially in densely settled rural areas.[36] As we predicted, increasing household size, which we argue is a better indicator of the population-density conditions experienced by the average child than country-level population density itself, was correlated with worsening child health. Family size is correlated with many different conditions, including maternal education level (which might explain in part the relationship we observed). Nevertheless, the negative relationship between human
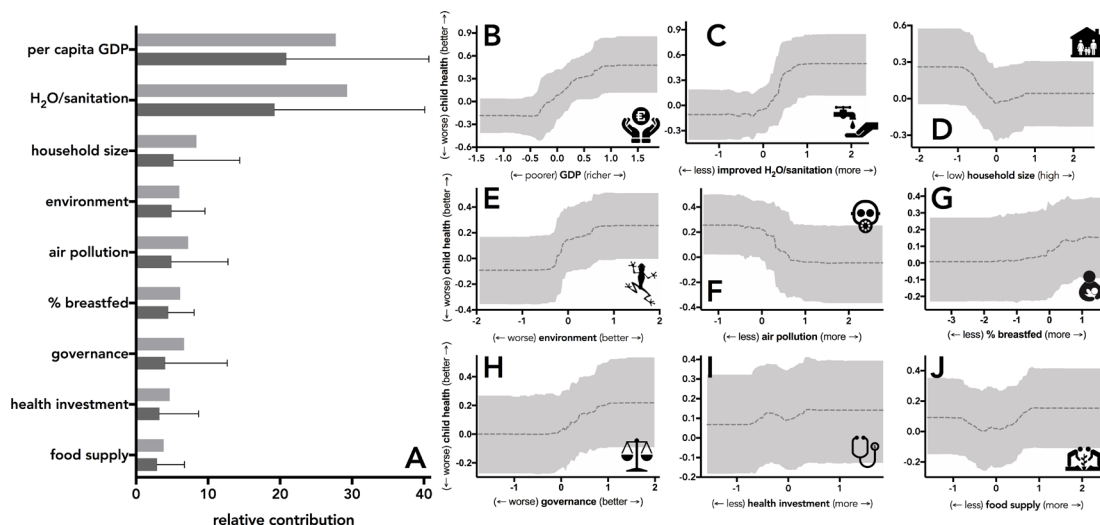
Figure 2 (A) The proportion of variance in child-health index explained by the socioeconomic and environmental variables. Light grey bars indicate the relative contribution of each variable to the variance in child health among countries from the full dataset, whereas the dark grey bars (and error bars) indicate the same results for the resampled boosted regression trees. Predicted child health (composite index) as a function of variation in (B) per capita wealth (gross domestic product (GDP)), (C) access to improved water and sanitation, (D) mean household size, (E) environmental performance, (F) air pollution, (G) % of children exclusively breast fed for the first 6 months of life, (H) governance quality, (I) per capita health investment and (J) per capita food supply (kcal/person/day).

population density and environmental performance in Africa,[29] and projections showing that many African nations will increase rapidly in population size over the coming decades,[9 11] raise concerns that increased population pressure will further compromise child health.

The relatively strong environmental signal we found is somewhat unexpected, given the anticipated lag between destructive actions like deforestation and pollution to be expressed in terms of human health.[61] In most regions of Africa, this result suggests that environmental degradation is possibly now already at a point where it is compromising food production, water or air quality or defence against infectious disease.[61–63] From a climate-change perspective, there is evidence that rainfall shocks associated with a more variable future climate at least temporarily increase stunting and the incidence of diarrhoea in children.[64]

Despite evidence that food security and gut health in infants and young children benefit greatly from breast

feeding,[23–25] the role of exclusive breastfeeding for the first 6 months of life was relatively weak in explaining variance in child health, although the correlation was in the hypothesised direction. Likewise, governance quality was also a weak predictor of child health, but again in the predicted direction. There is evidence that while globalisation and improved democracy reduce infant mortality,[33] the relationship might be weakened by the definition of governance we used here, which focuses on performance at the political level and not necessarily on the development and governance of health and education systems.[65 66]

The two weakest correlates of national-scale performance in child health were health investment and food supply. While past investments in health interventions such as immunisation and the provision of insecticide-treated bed nets have been instrumental[2] in countries such as Tanzania for reducing child mortality,[67] there is evidence that the International Monetary fund—which determines how governments allocate scarce resources to public-health interventions—has in fact decreased the protective effect of parents' education on child malnourishment, sanitation, healthcare access, immunisation and shelter, especially in rural areas[68] (via changed priorities in government spending and wage-bill ceilings that lead to fewer teachers and reduced teaching quality). While increases in health spending should benefit child health, external forces can instead dismantle existing government subsidies and subvert the progress towards improving overall health outcomes.[68] The weak, non-linear relationship between child health and food supply appeared to result from a combination of a direct beneficial effect on stunting and weaker variable effects on other metrics of

Table 3 Goodness-of-fit statistics for boosted regression trees using individual health metrics as the response variable

Health metric	trees	D^2	β_{cv}	SE β_{cv}
Stunting	26 300	50.5	49.0	15.9
Respiratory infection	51 200	58.7	72.7	8.6
Diarrhoeal disease	40 250	56.3	66.4	10.0
Infectious diseases	26 050	41.4	39.3	13.3
Injury	40 850	62.0	49.1	21.2

Shown is the final number of trees evaluated (trees), percentage of deviance explained (D^2), cross-validation correlation coefficient (β_{cv}) and its standard error calculated over all tree iterations (SE β_{cv}).
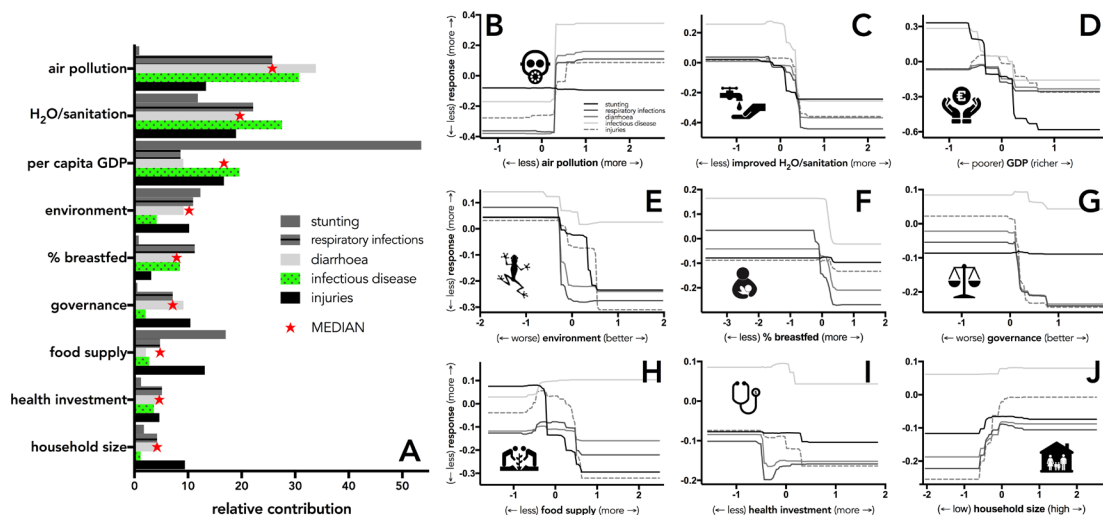
Figure 3 (A) The proportion of variance in each of the individual health indicators (stunting, respiratory infections, diarrhoeal disease, infectious disease, injury) explained by the socioeconomic and environmental variables. Predicted child-health indicators as a function of variation in (B) air pollution, (C) access to improved water and sanitation, (D) per capita wealth (gross domestic product (GDP)), (E) environmental performance, (F) % of children exclusively breast fed for the first 6 months of life, (G) governance quality, (H) per capita food supply (kcal/person/day), (I) per capita health investment and (J) household size.

child health. This non-linearity might also be explained partially by the substantial subnational spatial variation in child undernutrition across the continent.[2] While many regions of Africa have in the past experienced extensive food shortages that had long-lasting effects on child health, today it appears that other aspects of environment and socioeconomic policies are more important factors.

CONCLUSION

Clearly, the complex interaction of environmental and socioeconomic conditions that influence child-health outcomes in Africa makes generalisations regarding the prioritisation of interventions difficult, and the most effective, regionally specific interventions will necessarily be highly variable. Nonetheless, our broad-scale assessment indicates that on average, increasing the long-term outcomes of child health in Africa are potentially most achievable through relatively simple interventions such as improvement of water, air quality, and sanitation. Complex issues such as slowing the seemingly inevitable and rapid rise in population density[9] through culturally appropriate, ethical and supportive family planning,[69] and addressing the rising loss of biodiversity and overall environmental integrity, while more challenging, also cannot be ignored in planning for the future health of children in Africa.

Author affiliations
[1]Global Ecology, College of Science and Engineering, Flinders University, Adelaide, South Australia, Australia
[2]Department of Zoology, University of British Columbia, Vancouver, British Columbia, Canada
[3]Discipline of Paediatrics, University of Western Australia, Crawley, Western Australia, Australia
[4]Center on Food Security and the Environment, Stanford University, Stanford, California, USA
[5]Rand Corporation, Santa Monica, California, USA

ORCID iD
Corey J A Bradshaw http://orcid.org/0000-0002-5328-7741

REFERENCES
1 World Health Organization. *Children: reducing mortality fact sheet*. Geneva: World Health Organization, 2017.
2 Osgood-Zimmerman A, Millear AI, Stubbs RW, *et al*. Mapping child growth failure in Africa between 2000 and 2015. *Nature* 2018;555:41–7.
3 UNICEF, World Health Organization. *Countdown to 2030: tracking progress towards universal coverage for reproductive, newborn*

and child health: the 2017 report. Washington, DC: United Nations Children's Fund and the World Health Organization, 2017.

4 Kirigia JM, Muthuri RDK, Nabyonga-Orem J, *et al*. Counting the cost of child mortality in the world Health organization African region. *BMC Public Health* 2015;15:1103.

5 Heft-Neal S, Burney J, Bendavid E, *et al*. Robust relationship between air quality and infant mortality in Africa. *Nature* 2018;559:254–8.

6 Zhang Y, Bi P, Hiller JE. Climate change and disability-adjusted life years. *J Env Health* 2007;70:32–6.

7 Rachel C. Public health implications and risks for children and families resettled after exposure to armed conflict and displacement. *Scand J Pub Health* 2016;45:209–11.

8 Burke M, Heft-Neal S, Bendavid E. Sources of variation in under-5 mortality across sub-Saharan Africa: a spatial analysis. *Lancet Glob Health* 2016;4:e936–45.

9 Bradshaw CJA, Brook BW. Human population reduction is not a quick fix for environmental problems. *Proc Natl Acad Sci U S A* 2014;111:16610–5.

10 Gerland P, Raftery AE, Sevčíková H, *et al*. World population stabilization unlikely this century. *Science* 2014;346:234–7.

11 Bongaarts J, Casterline J. Fertility transition: is sub-Saharan Africa different? *Popul Dev Rev* 2013;38:153–68.

12 Ceballos G, Ehrlich PR, Dirzo R. Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proc Natl Acad Sci U S A* 2017;114:E6089–96.

13 Grace K, Frederick L, Brown ME, *et al*. Investigating important interactions between water and food security for child health in Burkina Faso. *Popul Environ* 2017;39:26–46.

14 Annim SK, Awusabo-Asare K, Amo-Adjei J. Household nucleation, dependency and child health outcomes in Ghana. *J Biosoc Sci* 2015;47:565–92.

15 Root G. Population density and spatial differentials in child mortality in Zimbabwe. *Soc Sci Med* 1997;44:413–21.

16 Anand A, Roy N. Transitioning toward sustainable development goals: the role of household environment in influencing child health in sub-Saharan Africa and South Asia using recent demographic health surveys. *Front Public Health* 2016;4.

17 Garenne M. Urbanisation and child health in resource poor settings with special reference to under-five mortality in Africa. *Arch Dis Child* 2010;95:464–8.

18 Owili P, Lien W-H, Muga M, *et al*. The associations between types of ambient PM2.5 and under-five and maternal mortality in Africa. *Int J Environ Res Public Health* 2017;14:359.

19 Hathi P, Haque S, Pant L, *et al*. Place and child health: the interaction of population density and sanitation in developing countries. *Demography* 2017;54:337–60.

20 Vrijheid M, Casas M, Gascon M, *et al*. Environmental pollutants and child health—a review of recent concerns. *Int J Hyg Environ Health* 2016;219:331–42.

21 Smith LC, Haddad L. Reducing child undernutrition: past drivers and priorities for the post-MDG era. *World Dev* 2015;68:180–204.

22 Victora CG, Bahl R, Barros AJD, *et al*. Breastfeeding in the 21st century: epidemiology, mechanisms, and lifelong effect. *The Lancet* 2016;387:475–90.

23 Rautava S. Early microbial contact, the breast milk microbiome and child health. *J Dev Orig Health Dis* 2016;7:5–14.

24 Baker P, Smith J, Salmon L, *et al*. Global trends and patterns of commercial milk-based formula sales: is an unprecedented infant and young child feeding transition underway? *Public Health Nutr* 2016;19:2540–50.

25 Salmon L. Food security for infants and young children: an opportunity for breastfeeding policy? *Int Breastfeed J* 2015;10:7.

26 Lovasi GS, Quinn JW, Neckerman KM, *et al*. Children living in areas with more street trees have lower prevalence of asthma. *J Epidemiol Community Health* 2008;62:647–9.

27 Bell JF, Wilson JS, Liu GC. Neighborhood greenness and 2-year changes in body mass index of children and youth. *Am J Prev Med* 2008;35:547–53.

28 Wolch JR, Byrne J, Newell JP. Urban green space, public health, and environmental justice: The challenge of making cities 'just green enough'. *Landsc Urban Plan* 2014;125:234–44.

29 Bradshaw CJA, Di Minin E. Socio-economic predictors of environmental performance among African nations. *Sci Rep* 2019;9:9306.

30 Chen Y, Lei X, Zhou L-A. Does raising family income cause better child health? Evidence from China. *Econ Dev Cult Change* 2017;65:495–520.

31 World Health Organization. *World health statistics 2017: monitoring health for the SDGs, sustainable development goals*. Geneva: World Health Organization, 2017.

32 Marmot M, Friel S, Bell R, *et al*. Closing the gap in a generation: health equity through action on the social determinants of health. *The Lancet* 2008;372:1661–9.

33 Welander A, Lyttkens CH, Nilsson T. Globalization, democracy, and child health in developing countries. *Soc Sci Med* 2015;136-137:52–63.

34 Wolfe BL, Behrman JR. Determinants of child mortality, health, and nutrition in a developing country. *J Dev Econ* 1982;11:163–93.

35 Lolekha S, Warachit B, Hutagalung Y, *et al*. Effect of climatic factors and population density on varicella zoster virus epidemiology within a tropical country. *Am J Trop Med Hyg* 2001;64:131–6.

36 Greiner KA, Li C, Kawachi I, *et al*. The relationships of social participation and community ratings to health and health behaviors in areas with high and low population density. *Soc Sci Med* 2004;59:2303–12.

37 R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2018.

38 Bradshaw CJA, Giam X, Sodhi NS. Evaluating the relative environmental impact of countries. *PLoS One* 2010;5:e10440.

39 van Donkelaar A, Martin RV, Brauer M, *et al*. Global estimates of fine particulate matter using a combined geophysical-statistical method with information from satellites, models, and monitors. *Environ Sci Technol* 2016;50:3762–72.

40 Global Administrative Areas. *GADM database of global administrative areas*. Version 2.8, 2015. www.gadm.org.

41 Bright EA, Rose AN, Landscan UML. *High-resolution global population data set*. 2016. Oak Ridge, Tennessee, USA: Oak Ridge National Laboratory, 2015.

42 Lin D, Hanscom L, Martindill J. *Working Guidebook to the National footprint accounts: 2016 edition*. Oakland, California: Global Ecological Footprint Network, 2016. footprintnetwork.org

43 Lindsey PA, Chapron G, Petracca LS, *et al*. Relative efforts of countries to conserve world's megafauna. *Glob Ecol Conserv* 2017;10:243–52.

44 Hansen MC, Potapov PV, Moore R, *et al*. High-Resolution global maps of 21st-century forest cover change. *Science* 2013;342:850–3.

45 Dieleman JL, Schneider MT, Haakenstad A, *et al*. Development assistance for health: past trends, associations, and the future of international financial flows for health. *The Lancet* 2016;387:2536–44.

46 Mo Ibrahim Foundation. *Ibrahim index of African governance. Detailed methodology*. London, United Kingdom: Mo Ibrahim Foundation, 2016.

47 Luck GW. A review of the relationships between human population density and biodiversity. *Biol Rev Camb Philos Soc* 2007;82:607–45.

48 Aaby P. Lessons for the past: third world evidence and the reinterpretation of developed world mortality declines. In: *Health transition review: historical epidemiology and the health transition*. , 1992: Supplement 2, 155–84.

49 Caldwell JC. *Theory of fertility decline*. New York: Academic Press, 1982.

50 Elith J, Leathwick JR, Hastie T. A working guide to boosted regression trees. *J Anim Ecology* 2008;77:802–13.

51 Hijmans RJ, Phillips S, Leathwick J. *Dismo: species distribution modeling. R package 601 version 1.1-4*, 2017. CRAN.R-project.org/package=dismo

52 Breiman L. Statistical modeling: the two cultures (with comments and a rejoinder by the author). *Stat Sci* 2001;16:199–231.

53 Breiman L, Friedman JH, Olshen RA, *et al*. *Classification and regression trees*. Belmont, California, USA: Wadsworth International Group, 1984.

54 Prowse TAA, Bradshaw CJA, Delean S, *et al*. An efficient protocol for the global sensitivity analysis of stochastic ecological models. *Ecosphere* 2016;7:e01238.

55 Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Stat* 2001;29:1189–232.

56 Bradshaw CJA, Brook BW. How to rank journals. *PLoS One* 2016;11:e0149852.

57 Lehmann G. Kappa sigma clipping. *Insight J* 2006. http://hdl.handle.net/1926/367

58 Mitchell RJ. Testing evolutionary and ecological hypotheses using path analysis and structural equation modelling. *Funct Ecol* 1992;6:123–9.

59 Bloom DE, Canning D, Sevilla J. The effect of health on economic growth: a production function approach. *World Dev* 2004;32:1–13.

60 Bloom DE, Canning D. Policy Forum: public health. The health and wealth of nations. *Science* 2000;287:1207–9.

61 Butler CD, CJA B, Gillespie S, *et al*. *Research priorities for the environment, agriculture and infectious diseases of poverty. Technical report of the TDR thematic reference group on environment, agriculture and infectious diseases of poverty*. Rome, Italy: World

Health Organization (WHO), 2013: 125Technical Report Series No. 976.

62 McMichael AJ, Woodruff RE, Hales S. Climate change and human health: present and future risks. *Lancet* 2006;367:859–69.

63 McMichael T, Montgomery H, Costello A. Health risks, present and future, from global climate change. *BMJ* 2012;344:e1359.

64 Rabassa M, Skoufias E, Jacoby H. Weather and child health in rural Nigeria. *J Afr Econ* 2014;23:464–92.

65 Cole MA, Elliott RJR, Zhang J. Corruption, governance and FDI location in China: a province-level analysis. *J Dev Stud* 2009;45:1494–512.

66 Moore M, Leavy J, Houtzager P, *et al. Polity qualities: how governance affects poverty*. London, United Kingdom: United Kingdom Department for International Development, 1999.

67 Afnan-Holmes H, Magoma M, John T, *et al*. Tanzania's countdown to 2015: an analysis of two decades of progress and gaps for reproductive, maternal, newborn, and child health, to inform priorities for post-2015. *Lancet Glob Health* 2015;3:e396–409.

68 Daoud A, Nosrati E, Reinsberg B, *et al*. Impact of international monetary fund programs on child health. *Proc Natl Acad Sci U S A* 2017;114:6492–7.

69 United Nations. Safe, effective family planning is key to 'empowering people, developing nations' – UN, 2017. Available: www.un.org/sustainabledevelopment/blog/2017/07/safe-effective-family-planning-is-key-to-empowering-people-developing-nations-un