RESEARCH ARTICLE

REVISED Mortality in rural coastal Kenya measured using the Kilifi Health and Demographic Surveillance System: a 16-year descriptive analysis [version 2; peer review: 3 approved, 1 approved with reservations]

Mark Otiende[1], Evasius Bauni[1], Amek Nyaguara[1], David Amadi[1], Christopher Nyundo[1], Emmanuel Tsory[1], David Walumbe[1], Michael Kinuthia[1], Norbert Kihuha[1], Michael Kahindi[1], Gideon Nyutu[1], Jennifer Moisi[1], Amare Deribew[1], Ambrose Agweyu[1], Kevin Marsh[2], Benjamin Tsofa[1], Philip Bejon[1,2], Christian Bottomley[3], Thomas N. Williams[1], J. Anthony G. Scott[1,3]

[1]Epidemiology and Demography, KEMRI-Wellcome Trust Research Programme, Kilifi, 80108, Kenya
[2]Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK
[3]Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, UK

v2 **First published:** 29 Nov 2021, **6**:327
https://doi.org/10.12688/wellcomeopenres.17307.1
Latest published: 17 Feb 2023, **6**:327
https://doi.org/10.12688/wellcomeopenres.17307.2

Abstract

Background: The Kilifi Health and Demographic Surveillance System (KHDSS) was established in 2000 to define the incidence and prevalence of local diseases and evaluate the impact of community-based interventions. KHDSS morbidity data have been reported comprehensively but mortality has not been described. This analysis describes mortality in the KHDSS over 16 years.
Methods: We calculated mortality rates from 2003–2018 in four intervals of equal duration and assessed differences in mortality across these intervals by age and sex. We calculated the period survival function and median survival using the Kaplan–Meier method and mean life expectancies using abridged life tables. We estimated trend and seasonality by decomposing a time series of monthly mortality rates. We used choropleth maps and random-effects Poisson regression to investigate geographical heterogeneity.
Results: Mortality declined by 36% overall between 2003–2018 and by 59% in children aged <5 years. Most of the decline occurred between 2003 and 2006. Among adults, the greatest decline (49%) was observed in those aged 15–54 years. Life expectancy at birth increased by 12 years. Females outlived males by 6 years. Seasonality was only evident in the 1–4 year age group in the first four years. Geographical

Open Peer Review

Approval Status ✓ ? ✓ ✓

	1	2	3	4
version 2 (revision) 17 Feb 2023			✓ view	✓ view
version 1 29 Nov 2021	✓ view	? view		

1. **Stephen Gordon**, Liverpool School of Tropical Medicine, Liverpool, UK
Malawi Liverpool Wellcome Programme, Blantyre, Malawi

2. **Momodou Jasseh**, Medical Research Council Unit The Gambia at LSHTM, Fajara, The Gambia

3. **Ian Cook**, University of Limpopo,

variation in mortality was ±10% of the median value and did not change over time.
Conclusions: Between 2003 and 2018, mortality among children and young adults has improved substantially. The steep decline in 2003–2006 followed by a much slower reduction thereafter suggests improvements in health and wellbeing have plateaued in the last 12 years. However, there is substantial inequality in mortality experience by geographical location.

Keywords
Child mortality, adult mortality, all-cause mortality, demography, demographic surveillance system, mortality trends, Kenya

Mankweng, South Africa

4. **Alain Vandormael**, Heidelberg University, Im Neuenheimer Feld, Heidelberg, Germany

Any reports and responses or comments on the article can be found at the end of the article.

KEMRI wellcome

This article is included in the KEMRI | Wellcome Trust gateway.

Corresponding author: Mark Otiende (MOtiende@kemri-wellcome.org)

Author roles: Otiende M: Data Curation, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Bauni E**: Conceptualization, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Supervision; **Nyaguara A**: Funding Acquisition, Investigation, Project Administration, Resources, Supervision, Writing – Review & Editing; **Amadi D**: Data Curation, Investigation, Project Administration, Software; **Nyundo C**: Investigation, Project Administration; **Tsory E**: Investigation; **Walumbe D**: Investigation, Project Administration, Resources; **Kinuthia M**: Software; **Kihuha N**: Software; **Kahindi M**: Software; **Nyutu G**: Investigation, Software; **Moisi J**: Investigation; **Deribew A**: Investigation; **Agweyu A**: Investigation; **Marsh K**: Conceptualization, Funding Acquisition, Investigation, Methodology, Project Administration, Writing – Review & Editing; **Tsofa B**: Funding Acquisition, Investigation; **Bejon P**: Funding Acquisition, Investigation, Project Administration, Writing – Review & Editing; **Bottomley C**: Formal Analysis, Investigation, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Williams TN**: Conceptualization, Funding Acquisition, Investigation, Methodology, Project Administration, Supervision, Writing – Review & Editing; **Scott JAG**: Conceptualization, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by Wellcome [OXF-COR03-2430] and [214320; to JAGS]
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Otiende M, Bauni E, Nyaguara A *et al.* **Mortality in rural coastal Kenya measured using the Kilifi Health and Demographic Surveillance System: a 16-year descriptive analysis [version 2; peer review: 3 approved, 1 approved with reservations]** Wellcome Open Research 2023, **6**:327 https://doi.org/10.12688/wellcomeopenres.17307.2

First published: 29 Nov 2021, **6**:327 https://doi.org/10.12688/wellcomeopenres.17307.1

REVISED **Amendments from Version 1**

We have updated the manuscript accordingly in response to the comments from the reviewers. Specifically, we have added additional information to Table 2 i.e., numerators and denominators for the mortality rates. We have conducted additional analyses to check whether the difference in mortality rates between periods could be due to random month-month fluctuations. All these have been described in the methods and reported in the results.

Any further responses from the reviewers can be found at the end of the article

Introduction

A majority of low- and middle-income countries (LMICs), especially in sub-Saharan Africa (sSA), lack comprehensive civil registration and vital statistics' systems (CRVS) necessary for monitoring mortality[1,2]. Tracking the progress in child and adult survival, therefore, relies on alternative data sources such as demographic and health surveys (DHS), national population censuses, and health and demographic surveillance systems (HDSSs). HDSSs are designed to monitor a small subpopulation of a nation in a defined geographical area and are a commonly used resource for health and demographic research in LMICs.

The Kilifi Health and Demographic Surveillance System (KHDSS) was established in 2000 by the KEMRI–Wellcome Trust Research Programme (KWTRP) to monitor mortality and morbidity caused by common diseases and to provide a sampling frame for epidemiological studies[3]. The surveillance area was selected to capture at least 80% of patients admitted to the Kilifi County Hospital (KCH) and over the subsequent two decades, the platform has been used to describe morbidity in children and adults. This includes incidence of malaria[4], pneumonia[5], lower respiratory tract infections[6], rotavirus[7], malnutrition[8,9], sickle cell disease[10], epilepsy[11] as well as the burden of mental health problems[12], pregnancy-related disorders and chronic diseases that contribute substantially to overall mortality[13,14]. The KHDSS has also been used to evaluate the impact of new community-based interventions, such as vaccines and bed net use[5,15–19].

Morbidity data have been reported comprehensively in Kilifi but mortality in this population has not been systematically described. In this paper, we describe mortality in children and adults over a 16-year period and analyse deaths by age, sex, season, geographical location, and temporal trend.

Methods

Data source and setting

The KHDSS surveillance area, which is located in Kilifi County, within the former Coast Province, is divided into 15 government administrative regions called locations, comprising a total of 891 km^2 (Figure S1, Extended data[20]). An initial census and mapping of the surveillance area was conducted in 2000 and was found to contain 189,148 residents in 20,978 households. Subsequently, the population has been continuously monitored for births, pregnancies, deaths, and migration events in re-enumeration rounds occurring approximately every 4 months, and the mapping was updated in 2017. At the end of 2018, there were 299,471 residents living in 41,536 households. From 2008, deaths registered in the KHDSS have also been investigated for the cause of death by verbal autopsy which has been reported separately[21]. KCH is located at the geographical centre of the KHDSS area and, during the study period, it was the only government facility offering inpatient care for the KHDSS population. A small number of private hospitals and lower-level facilities have a few inpatient beds.

The concept of the KHDSS is based on the INDEPTH (International Network for the Demographic Evaluation of Populations and Their Health) data model. Demographic and health data are collected at four points of contact: at re-enumeration when community interviewers make household visits to update the population register; at the inpatient wards of KCH where medical staff record patient history, clinical examination and outcome (death or discharge); at the maternity ward of KCH where staff record births and perinatal deaths; and in 34 vaccination clinics distributed across the surveillance area which collected data on childhood vaccination between 2008 and 2018. The eligibility for inclusion, the variables routinely measured, the structure of the KHDSS databases and the population structure have all been described previously[3].

Initially, data collection during household visits was paper based but switched to electronic data collection using tablets in 2016. The tablets are loaded daily with the most recent copy of the residents' database and, after data collection, they are returned to the research unit where a two-way synchronization with the master database is performed. All other data collection points are linked in real time to the master database that has been specified using MySQL.

At re-enumeration rounds, information on all household members is sought from a single informant, usually a member of the household. If all household members are unavailable during the visit, information is obtained from neighbouring households. All field staff are debriefed on the quality of data collected after each enumeration cycle and re-trained where needed. The data collection applications are programmed with skip patterns and consistency checks to ensure mandatory information is collected. Additionally, within the database, there are built-in checks for missing or duplicated data.

To explore the accuracy of age data at the first census and among all new in-migrants, we calculated Whipple's Index[22]. Whipple's index measures the tendency for individuals to inaccurately report their age in rounded numbers, usually ending in 0 and 5, resulting in age heaping.

Statistical analysis

The analysis period, from 1 Jan 2003 to 31 Dec 2018, was stratified into four non-overlapping periods each lasting 4 years. We excluded 2000–2002 because of changes in the re-enumeration

protocols designed to increase the ascertainment of deaths in neonates during these years. We used survival analysis and routine demographic life table methods to calculate mortality rates and life expectancy and examined seasonality, short- and long-term trends over the 16 years.

Age–sex mortality profile. The mortality rate was calculated as the number of deaths divided by person-years of observation (PYO). Entry to risk begins at the latest of birth, in-migration or study start date. Exit from risk is at the earliest of study end date, out-migration or death. If an out-migration is followed by an in-migration, the period between the out-migration and in-migration is excluded from the risk period to avoid survivor bias. The total PYO was computed for different age groups, sex, and locations.

For children aged less than five years, we have also calculated conventional mortality ratios where the number of deaths within a specific age group in a given time period is divided by the number of live births occurring during the same time period. Mortality ratios are commonly used in settings where risk time cannot be quantified. They can be confounded by varying birth rates as the deaths in the numerator are not always drawn from the denominator population.

Survival and life expectancy. We used two methods to estimate life expectancy: the period life table method which calculates the *mean* life expectancy at birth and the Kaplan–Meier (KM) survival method which calculates the *median* age at death. The main difference between the methods is in the age intervals used; the life table method computes survival probabilities within pre-define age intervals, *e.g.* 5-year intervals, whereas the KM method computes survival probabilities whenever there is a death in the cohort making the KM intervals smaller and of variable length[23].

For purposes of comparison with other analyses, we also generated abridged life tables using data structured according to analytic methods developed by the Multi-centre Analysis of the Dynamics of Internal Migration and Health (MADIMAH) which was a working group within INDEPTH. In the MADIMAH method[24,25], the definition of risk time considers the time between out-migration and a subsequent in-migration. If the difference is less than 180 days, this time is included in the risk period which increases the person-years of observation resulting in lower estimates of mortality rates.

Seasonality and trend. We first assessed seasonality and long-term temporal trends for each age group by graphically reviewing a time series of monthly mortality rates. We then estimated trend and seasonality based on an STL (Seasonal and Trend using LOESS) decomposition and identified months with the highest and lowest mortality rates from the seasonal component. We investigated whether the difference in mortality rates between analysis periods could be attributed to random month-month fluctuations by fitting a negative binomial model with yearly counts as the outcome and period as a categorical explanatory variable. To test for seasonality, we fitted a model that included month, period and the interaction to monthly mortality counts.

Geographical heterogeneity in survival and mortality over time. We produced choropleth maps for overall and age-specific mortality rates in the four 4-year periods to investigate the geographical variation of mortality in the administrative locations over time. For overall mortality, we accounted for temporal differences in the population age–sex structure by direct standardization against the 2011 KHDSS age-sex structure. All the maps were created at the administrative location level.

We used the quantile method to create five mortality rate classes for map reading. This method places equal numbers of data units (death rates) in each class resulting in classes centred on the median death rate. For each age group, the quintiles are derived from the entire mortality rate range between 2003 and 2018 and the resulting classification is applied across each of the 4-year-period maps for that age group. The quantile method, though simple, has been shown to be the optimal classification method for displaying geographically varying data in series in a map reading experiment[26].

To assess geographical variation in mortality, for each period, we fitted a multi-level Poisson regression model adjusting for sex and age in which location was included as a random effect and used the variance of the random effect to quantify heterogeneity. We tested for between-location variation within each period using the likelihood ratio test and tested for temporal variation in mortality rates between the 2003–2006 period and each of the subsequent periods using a z-test. We also calculated the median age at death for each of the 15 administrative locations in the 4-year periods and assessed variation in life expectancy by location and time.

All analyses were conducted using STATA/IC version 15.1 (StataCorp College Station, Texas, USA, RRID:SCR_012763) and R version 4.1.0 (RRID:SCR_001905)[27].

Results

Age–sex profile

The cohort consisted of 699,841 individuals of whom 125,587 (18%) were followed from birth. In total, we observed 22,207 deaths in 3,897,529 person-years. More than 95% of the information on residence and vital status was collected from respondents living in the same household. Females contributed 53% of the total PYO and 48% of deaths (Table 1). There was no indication of age heaping or misspecification of sex (Table S1, Extended data[20]).

Mortality was highest in the first time period (2003–2006) compared with the other time periods for all ages except adults aged >54 years (Figure S2a, Extended data[20]). The mortality rate in children aged <5 years declined by 44% between the periods 2003–2006 and 2007–2010, from 13.1 to 7.4 deaths per 1000 PYO, and continued to decline more slowly, reaching 5.4 in 2015–2018; we observed a similar pattern in all the

Table 1. Distribution of births, deaths, and person-years by sub-period.

	2003–06		2007–10		2011–14		2015–18		Total	
	N	%	N	%	N	%	N	%	N	%
Live births	30572		31912		32994		30109		125587	
PYOs										
Males	397150	47.0	441785	46.8	485654	47.1	508679	47.2	1833268	47.0
Females	448295	53.0	501306	53.2	546412	52.9	568248	52.8	2064261	53.0
0–4 y	158147	18.7	171108	18.1	177401	17.2	168683	15.7	675339	17.3
5–14 y	263370	31.2	296311	31.4	323398	31.3	335392	31.1	1218471	31.3
15–54 y	365673	43. 3	407115	43.2	452325	43.8	479750	44.5	1704863	43.7
55–74 y	49807	5.9	57529	6.1	66533	6.4	77251	7.2	251120	6.4
≥75 y	8449	1.0	11029	1.2	12408	1.2	15851	1.5	47737	1.2
Deaths										
Males	3232	50.2	2675	51.0	2783	52.7	2763	52.7	11453	51.6
Females	3204	49.8	2574	49.0	2500	47.3	2476	47.3	10754	48.4
U5 deaths										
0–28 d	765	36.8	488	38.7	437	38.6	393	43.1	2083	38.7
29–365 d	548	26.4	366	29.0	323	28.5	245	26.9	1482	27.5
1–4 y	763	36.8	407	32.3	372	32.9	273	30.0	1815	33.7
5–14 years deaths										
5–14 y	370		248		281		275		1174	
≥15 years deaths										
15–54 y	2089	52.4	1517	40.6	1462	37.8	1403	34.6	6471	41.3
55–74 y	1254	31.4	1342	35.9	1496	38.7	1450	35.8	5542	35.4
≥75 y	647	16.2	881	23.6	912	23.6	1200	29.6	3640	23.3

finer age-strata of children aged <5 years (Table 2). Results from fitting a negative binomial regression model to yearly counts confirmed that mortality rates were significantly lower in the latter periods, for all age groups <55 years, compared to the first time period (Table S2, Extended data[20]). Mortality ratios, per 1000 live births, for children aged <5 years, are presented in Table S3 (Extended data[20]). Over the whole 16-year period mortality rates were lowest in children aged 5–14 years (Figure S2b, Extended data[20]). In this age group, mortality declined from 1.4/1000 PYO in 2003–2006 to 0.8/1000 PYO in 2007–2010 (Table 2) and varied little thereafter. In adults, the steepest period-to-period decline (35%) was seen in the age group 15–54 years between the periods 2003–2006 and 2007–2010. Mortality changed very little over time for those aged ≥75 years. Overall, mortality was higher in males than females at all ages with differences being greater in adulthood than in childhood.

Survival

Median survival increased from 65 in 2003 to 77 in 2018. Females had higher survival throughout the study period and died, on average, 6 years later than males (Figure 1). The survival functions for males and females are similar until approximately 40 years of age after which females experience better survival (Figure 2 A–D). The improvement in survival over time was also greater in females than males (Figure 2 E–F).

Mean life expectancy, estimated using the abridged life table method, is 4–8 years lower than Kaplan–Meier median survival estimates (Table S4, Extended data[20]). Tables S5a to S5h (Extended data[20]) show the sex-specific abridged life tables for each of the 4-year periods. The life expectancy estimates calculated using the MADIMAH technique differ little from those estimated using the standard technique (Tables S6a to S6h, Extended data[20]).

Table 2. **Mortality rates per 1000 PYO by age stratum and time period.**

Age	2003–06			2007–10			2011–14			2015–18		
	Deaths	PYO	Rate	Deaths	PYO	Rate	Deaths	PYO	Rate	Deaths	PYO	Rate
0–28 days												
Total	765	2314.1	330.6	488	2430.1	200.8	437	2501.3	174.7	393	2276.9	172.6
Female	324	1144.9	283.0	208	1204.45	172.7	213	1229.12	173.3	159	1123.25	141.6
Male	441	1169.2	377.2	280	1225.6	228.5	224	1272.1	176.1	234	1153.6	202.8
29–365 days												
Total	548	29255.6	18.7	366	30807.6	11.9	323	31867.0	10.1	245	29918.3	8.2
Female	277	14518.9	19.1	176	15260.6	11.5	142	15728.2	9.0	123	14748.8	8.3
Male	271	14736.7	18.4	190	15547.0	12.2	181	16138.8	11.2	122	15169.5	8.0
<1 year												
Total	1313	31569.7	41.6	854	33237.6	25.7	760	34368.3	22.1	638	32195.2	19.8
Female	601	15663.7	38.4	384	16465.0	23.3	355	16957.4	20.9	282	15872.0	17.8
Male	712	15905.9	44.8	470	16772.6	28.0	405	17410.9	23.3	356	16323.2	21.8
1–4 years												
Total	763	126577.7	6.0	407	137870.1	3.0	372	143032.6	2.6	273	136487.8	2.0
Female	370	62814.0	5.9	183	68628.1	2.7	174	70801.2	2.5	112	67123.9	1.7
Male	393	63763.7	6.2	224	69242.0	3.2	198	72231.3	2.7	161	69363.9	2.3
0–4 years												
Total	2076	158147.3	13.1	1261	171107.8	7.4	1132	177400.8	6.4	911	168683.0	5.4
Female	971	78477.7	12.4	567	85093.1	6.7	529	87758.6	6.0	394	82996.0	4.7
Male	1105	79669.6	13.9	694	86014.6	8.1	603	89642.2	6.7	517	85687.0	6.0
5–14 years												
Total	370	263369.6	1.4	248	296311.0	0.8	281	323398.4	0.9	275	335392.0	0.8
Female	154	130156.8	1.2	110	147025.9	0.7	122	160082.3	0.8	114	165919.8	0.7
Male	216	133212.8	1.6	138	149285.1	0.9	159	163316.1	1.0	161	169472.2	1.0
15–54 years												
Total	2089	365673.2	5.7	1517	407114.5	3.7	1462	452325.5	3.2	1403	479750.2	2.9
Female	1164	206433.0	5.6	823	229603.4	3.6	743	251863.9	3.0	710	262419.3	2.7
Male	925	159240.3	5.8	694	177511.1	3.9	719	200461.6	3.6	693	217330.9	3.2
55–74 years												
Total	1254	49806.6	25.2	1342	57529.4	23.3	1496	66533.0	22.5	1450	77250.7	18.8
Female	608	28813.3	21.1	648	33535.7	19.3	692	39567.1	17.5	642	47135.0	13.6
Male	646	20993.3	30.8	694	23993.7	28.9	804	26965.9	29.8	808	30115.6	26.8
≥75 years												
Total	647	8449.0	76.6	881	11028.6	79.9	912	12408.1	73.5	1200	15850.9	75.7
Female	307	4414.7	69.5	426	6047.9	70.4	414	7139.8	58.0	616	9777.7	63.0
Male	340	4034.3	84.3	455	4980.6	91.4	498	5268.2	94.5	584	6073.2	96.2

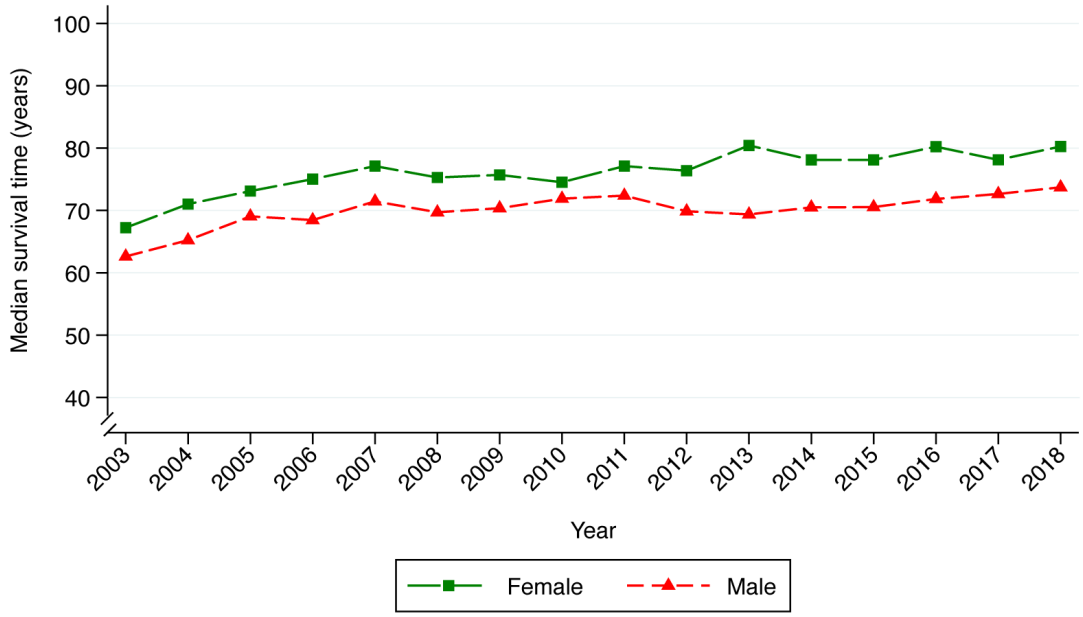

Figure 1. Median age at death based on the Kaplan–Meier method. Plot color represents sex with green and red indicating life expectancy in females and males respectively.

Seasonality and trend

Figure 3a shows the age-specific mortality rates by calendar month for each of the analysis periods. A seasonal pattern appears only in children aged 1–4 years in the first period. Table S7 (Extended data[20]) shows the estimated high and low mortality months. The months with the highest mortality for neonates, children aged 29–365 days and those aged 1–4 years were February, June, and July respectively. For adults aged 15–74 and ≥75 years, high mortality months were June and August, respectively. The interaction between period and seasonality was significant in children aged 1-4 years only (Table S2, Extended data[20]).

Figure 3b shows the trend in monthly mortality in three broad age groups. For children aged under five years, there was a steep decline in mortality from 2003 to 2008 followed by a gentle decline in the subsequent years. Within the same period, there was a similar pattern of decline in mortality rates in the older age groups. Figures S3a to S3g (Extended data[20]) show the 16-year mortality trends in finer age strata for both children and adults.

Geographical heterogeneity in survival and mortality

The largest span in median survival between locations was 7 years (highest in Gede, lowest in Junju and Tezo) in the period 2003–06 (Table S8, Extended data[20]). The span between highest and lowest location was reduced by 1 year in each of the subsequent 4-year study periods. Gains in life expectancy varied by location with one location (Sokoke) gaining 10 years after the first period and another location (Matsangoni) remaining unchanged. Changes in life expectancy between the periods 2003–2006 and 2007–2010 were greater in magnitude in comparison to any other two periods.

Figure 4 shows the geographical distribution of the age-standardized mortality rates for each population in the 15 locations over time. Overall mortality declined in all the locations but for those aged ≥55 years the decline was slow in some locations. (Figure S4a and S4b, Extended data[20]). Table 3 summarizes the geographical variation in mortality. After accounting for age and sex, variation was greatest in the period 2003–2006 with 95% of the location-specific rates lying within 16% of the median rate in this period. In the periods 2007–2010 and 2011–2014, we observed a decrease in variation with 95% of the location-specific rates lying within 9% of the median rate in both periods followed by an increase in 2015–2018 at 14%. Between-location variation was significant in the 2003–2006 and 2015–2018 periods. There were no significant differences in geographical variation in mortality between 2003–2006 and any of the subsequent periods. The output from the random effects Poisson regression model is summarized in Table S9 (Extended data[20]).

Discussion

In the Kilifi Health and Demographic Surveillance System, overall mortality rates declined steeply in the first four years of the study period in all age–sex groups, except in older adults, and then declined much more slowly in the subsequent

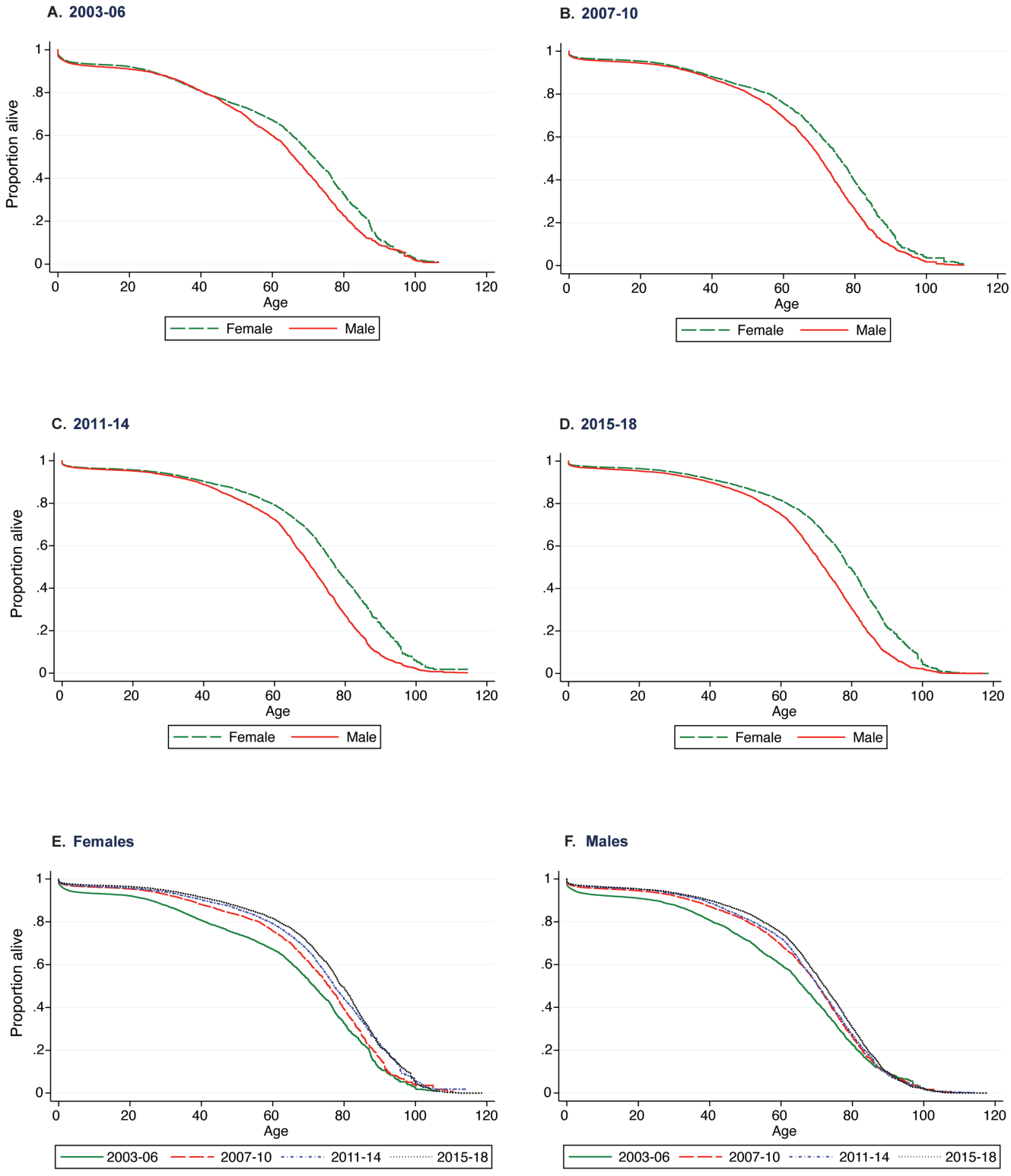

Figure 2. Kaplan–Meier survival curves for the KHDSS population by time period and sex. Plots **A**–**F** show the sex-specific survival curves and plots **E** and **F** show the period-specific survival curves.

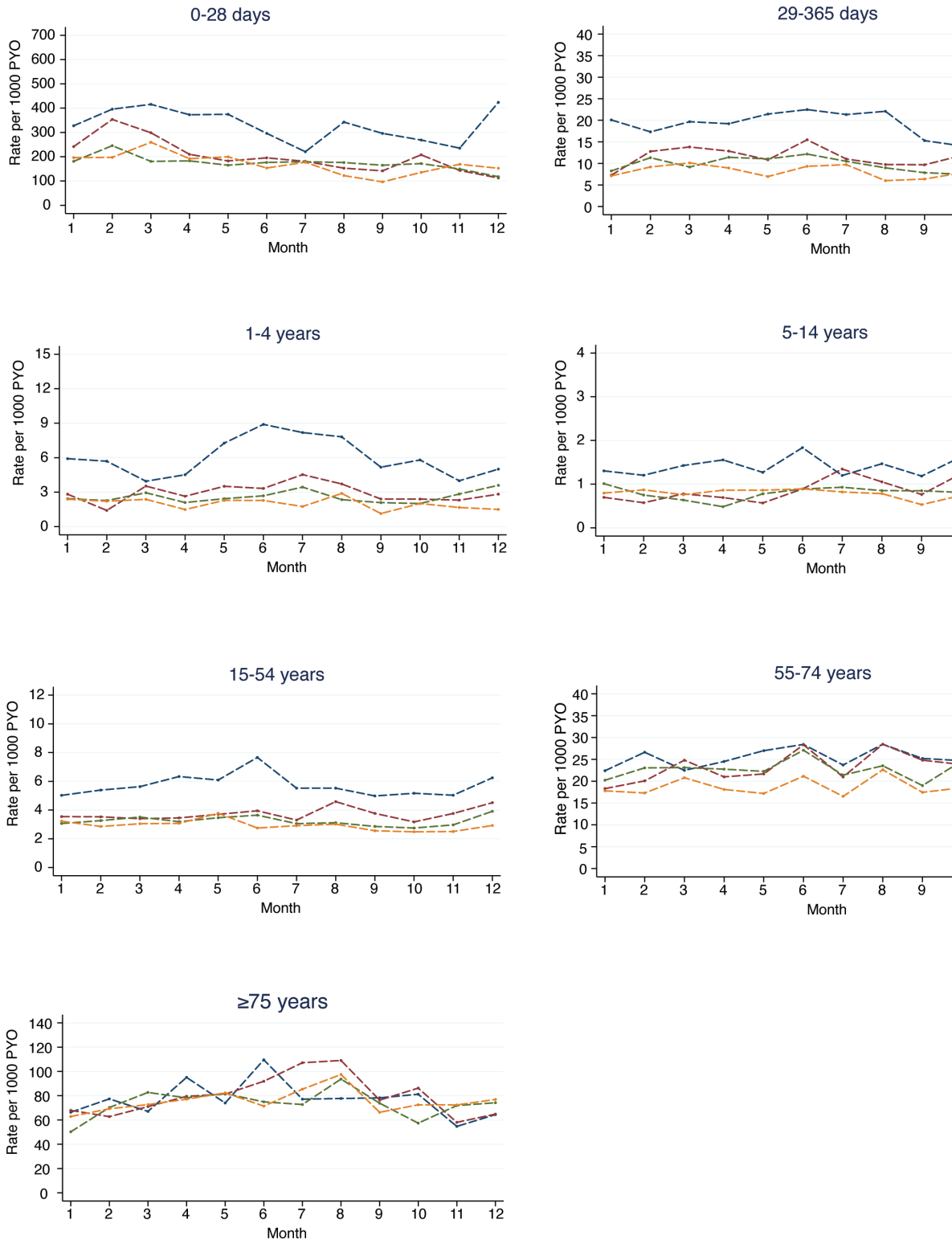

Figure 3a. Seasonal mortality rates (by calendar month) by age groups and time period.

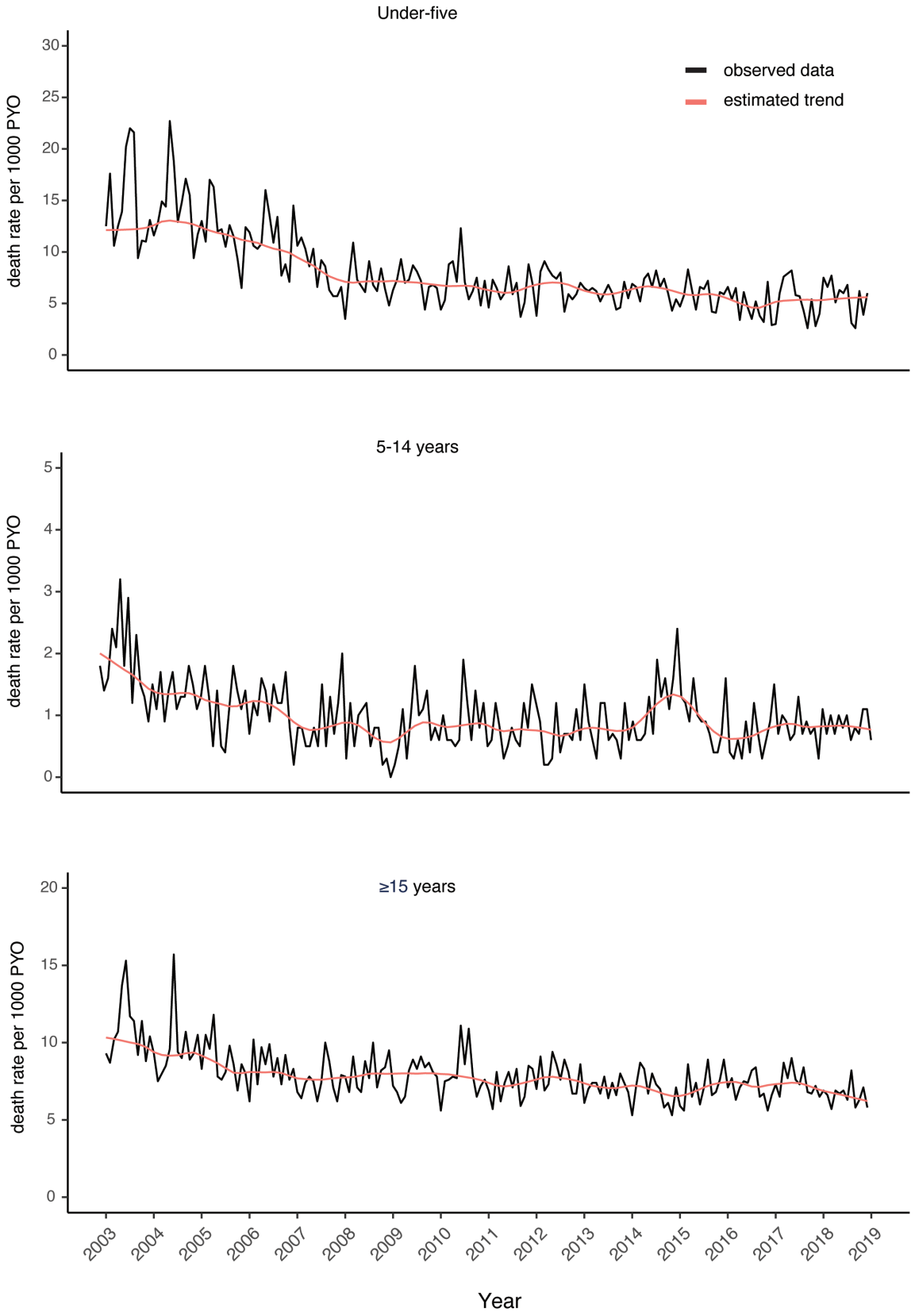

Figure 3b. Monthly mortality rates by age over the entire 16-year period. The black line represents the observed rates while the red line is the estimated trend.

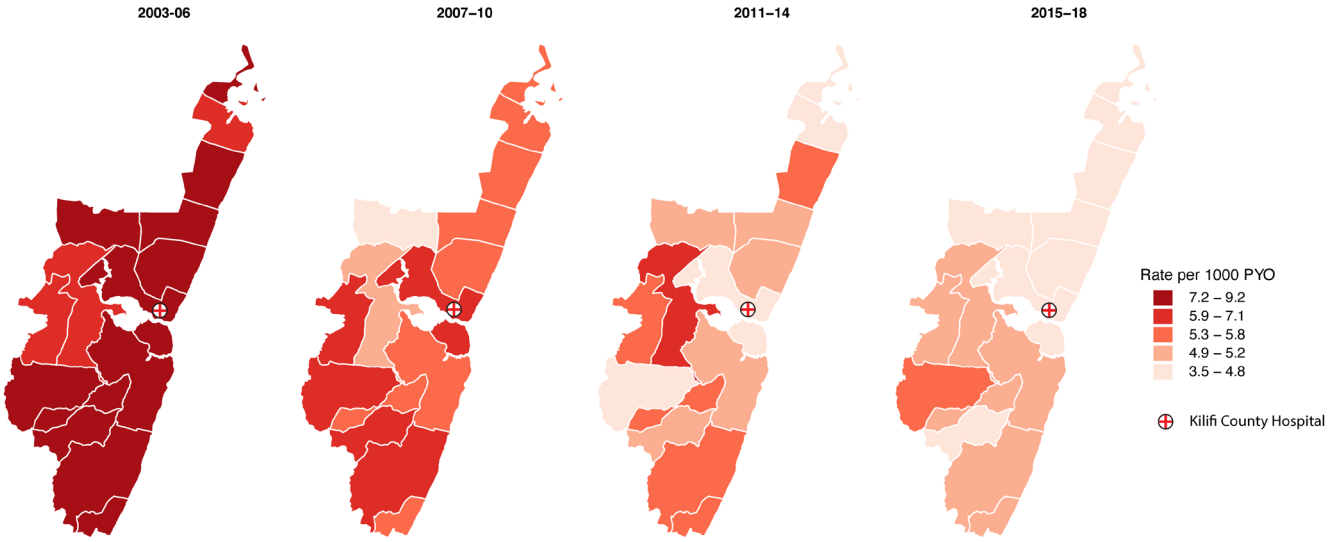

Figure 4. Age-standardized mortality rates by location and time period. Mortality rates were standardized to the KHDSS population structure in 2011. Darker shades indicate higher mortality. (Source: Own elaboration using shapefiles and data from KEMRI-Wellcome Trust Research Programme).

Table 3. Variation of location-specific $\log_e$ mortality rates (per 1000 PYO) adjusted by age and sex.

Period	SD of log rates[a]	95% normal approximation[b]	P-value (LR test of between-location variation)[c]	Differences between SDs[d]	P-value (z-test for equality of SD)[e]
2003–06	0.0778	1.16	<0.001		
2007–10	0.0444	1.09	0.072	0.033	0.267
2011–14	0.0433	1.09	0.067	0.034	0.246
2015–18	0.0670	1.14	0.002	0.011	0.715
2003–18	0.0498	1.10	<0.001		

Brief description of the analysis: SD of log rates are derived from a random effect Poisson regression model with administrative location included as the random effect. Geographical variation is quantified via SD of log rates, which is the square root of the variance of the random effect.

[a] Standard deviation - geographical (based on 15 administrative locations) variation in mortality rates. Calculated as the square root of the variance of the random effect.

[b] Approximation of the interval, about the overall median rate, that contains 95% of location-specific rates. Based on the normal distribution assumption that 95% of observations fall within 1.96 standard deviations of the median. Calculated as exp(SD of log rates *1.96). An example of the interpretation; in the period 2003–06, 95% of location-specific rates were within 16% of the median rate

[c] Likelihood ratio test comparing the fitted random effects model with that of a conventional Poisson model that does not include the random effect. Null hypothesis: no between-location variation in mortality rates

[d] Difference in SD between 2003-06 and subsequent periods

[e] P-value from test of equality of SDs between 2003–06 and subsequent periods

12 years. Neonatal and under-5 years mortality rates declined by 48% and 59%, respectively, between the first period and the last. The mortality reduction observed between the first and subsequent 4-year time periods was greatest among those aged 1–4 years. Median survival was greater in women, by 6 years compared with men, and increased in both sexes by approximately 12 years during the study period. Seasonal effects on mortality were only evident in children aged 1–4 years and

only in the first 4-year period. Finally, location-specific mortality varied from the median value by ±10%, which represents an important inequality. There was no evidence that this variation has improved over time.

Adult mortality in LMICs has received little attention in recent decades but from available estimates, adult mortality in East Africa either stagnated or declined between 2003 and 2010[28–30]. The trends coincide with earlier patterns of HIV incidence but these data are not sufficient to provide a detailed description of mortality levels and age patterns at national and sub-national levels, which hinders comparison. Nonetheless, adult mortality has been characterized by higher rates in females between 15–34 years because of deaths related to childbirth and possibly an earlier age of infection by HIV[31] (The Gap Report 2014). We observed this phenomenon in the first 4 years of the study period, but it was later reversed with men being at a greater risk of dying (Figure S5, Extended data[20]). This pattern is consistent with mortality sex ratios before and after the rapid initiation of HIV care and antiretroviral therapy in LMICs between 2003 and 2018 with evidence showing that women have benefitted more from the expansion of HIV treatment programs than men[32,33]. In Kilifi, the prevalence of HIV infection among women attending antenatal at KCH ranged between 3.8–4.4% between 2005–2009 and 2.0–3.7% between 2010–2016 with a clear decline from 2010[5].

Sub-national variation in child mortality, which is driven in part by inequitable distribution of health services and interventions, is a common observation across SSA[34–38]. In Kilifi, the magnitude of the overall variation can be understood as meaning that one location (5% of all locations) experienced a mortality rate that lies beyond ±10% of the average mortality rate. The magnitude of variation did not change significantly over time, which suggests sustained inequitable distribution in public health services and access to healthcare if we consider geographical variation in mortality as the measure of equity.

The decline in child mortality in Kilifi is consistent with independent observations over an extended period (Table S10, Extended data[20]). DHS data from Coast Province, which included Kilifi County, showed a reduction in child deaths (<5 years), per 1000 live births, from 116 to 57 between1993 and 2013. Where the two studies overlap in time (2004–2013), the DHS mortality ratios are higher than the KHDSS estimates; for example, the infant mortality ratios (IMR) are 31 in KHDSS and 44 in the Coastal DHS whilst the under-5 years mortality ratios (U5MR) are 46 and 57, respectively. The IMR estimated from the national census in 2009 was 42 per 1000 live births in Kilifi County for the 12 months preceding the census; the equivalent figure from KHDSS in 2008 was 25. For the U5MR the national census and KHDSS estimates were 57 and 38, respectively, per 1000 live births. Both the DHS and census rely on cross-sectional surveys and recall methods for death ascertainment whereas the KHDSS measures mortality directly from a cohort. Paradoxically, the methods that are dependent on potentially unreliable recall provide higher estimates of the number of deaths. The differences between the two methods are more likely to be driven by different definitions of their target populations. DHS and census methods capture all residents observed at one point in time within the geographical locale; some of these may not meet the residence requirement of the HDSS cohort. These requirements include that they are, or intend to be, resident in this household for at least 3 months. Furthermore, the DHS covers the whole of Coast Province, which has four counties in addition to Kilifi and, even within Kilifi, the KHDSS is a sub-population (approximately 40%) of the county. The overall numbers of child mortality from the HDSS, DHS and census datasets may be different but the trends are consistent with each other and are also consistent with an analysis of multiple disparate datasets that shows a decline in child mortality beginning as far back as 1965[34].

This descriptive analysis lays out the baseline trends in mortality in Kilifi over time. Several additional data sources may help to explore the underlying causes of these trends and geographic patterns. Within KHDSS there are data on the changing morbidity experience of residents from hospital records over the same period; for example, the incidence of admission to hospital with malaria declined sharply between 2003–2006[39]. Similar declines in LMICs have largely been attributed to reductions in malaria transmission following a high coverage of control measures[40]. The declining trend in malaria admissions in Kilifi is similar to the all-cause child mortality trend we have observed. Previous studies have reported a substantial indirect contribution of malaria to all-cause child mortality[41–43]. A study in this population found that malaria infection strongly predisposes individuals to bacteraemia, a major cause of childhood death, and could account for at least 50% of all bacteraemia cases in children[44]. The geographical heterogeneity in mortality (Figure 4) appears similar to the geographical heterogeneity in malaria[45], as has been observed with national heterogeneity in malaria and mortality[46]. In Kilifi, we observed a significant reduction in childhood morbidity after the introduction of *Haemophilus influenzae* type b conjugate vaccine in 2001[47], 10-valent pneumococcal conjugate vaccine in 2011[5,15] and rotavirus vaccine in 2014[16] but their effects on mortality have not yet been established. Finally, verbal autopsy data are also available to explore changes in causes of death over time[21]. The integration of these sources of data is beyond the scope of the present analysis.

Between 2003 and 2018, Kenya has experienced average GDP growth of 5.24% per annum and a population growth of 2.63% per annum (World Bank National Accounts Data). The area of KHDSS is typical of much of sub-Saharan Africa, being largely rural with a central town of approximately 43,000 residents (14% of the KHDSS). The detailed surveillance reported here over 16 years in KHDSS illustrates a clear improvement in mortality rates at all ages below 75 years but the pace of improvement has declined markedly over time and the geographical distribution of mortality rates is not

homogeneous. These findings highlight opportunities for intervention and improvement across a wide range of health, social and economic domains.

List of abbreviations

LMICs: Low- and Middle-Income Countries

sSA: sub-Saharan Africa

CRVS: Civil Registration and Vital Statistics Systems

DHS: Demographic and Health Surveys

HDSSs: Health and Demographic Surveillance Systems

KWTRP: KEMRI-Wellcome Trust Research Programme

KHDSS: Kilifi Health and Demographic Surveillance System

KCH: Kilifi County Hospital

INDEPTH: International Network for the Demographic Evaluation of Populations and Their Health

MySQL: Structured query language

PYO: Person years of observation

KM: Kaplan-Meier

MADIMAH: Multi-centre Analysis of the Dynamics of Internal Migration and Health

STL: Seasonal and Trend using LOESS

HIV: Human immunodeficiency virus

IMR: Infant mortality rate

U5MR: Under-five mortality rate

GDP: Gross domestic product

Declarations

Ethics approval and consent to participate

Individual verbal consent is sought at the household level using a specific informed consent form. Sensitization sessions through community stakeholders, including a hybrid community advisory board network comprising over 200 elected community representatives drawn from the KHDSS area. The analysis was approved by the Ethical Review Committee of the Kenya Medical Research Institute (approval number: KEMRI/SERU/CGMR-C/007/3057).

Data availability

Underlying data

Harvard Dataverse: Replication Data for: Mortality in rural coastal Kenya measured using the Kilifi Health and Demographic Surveillance System: a 16-year descriptive analysis. https://doi.org/10.7910/DVN/DDX7OK[20]

Underlying individual data include geo-located residence, date of birth and migration data and hence would be high risk for identifiability. This project contains the following intermediary data:

- Master ASMR - aggregated data with age specific mortality rates

Extended data

Harvard Dataverse: Replication Data for: Mortality in rural coastal Kenya measured using the Kilifi Health and Demographic Surveillance System: a 16-year descriptive analysis. https://doi.org/10.7910/DVN/DDX7OK[20]

This project contains the following extended data:

- Supplementary information - supplementary tables and figures

Author contributions

Conceptualization: JAGS, KM, TNW and EB. Data collection and preparation: MO, EB, AN, DA, CN, ET, DW, MW, NK, MK, GN. Analysis: MO, CB, JAGS Interpretation: MO, JAGS, CB, AN, AD, JM, AA, BT, PB, KM, JM, EB, TNW. Drafting the manuscript: MO, JAGS, CB, PB. Resources and funding acquisition: PB, JAGS, TNW, KM, BT, AN. All authors critically reviewed the article and approved the final version for submission.

Acknowledgements

We gratefully thank the residents of Kilifi who have participated in the surveillance activities of the KHDSS. We acknowledge the tremendous work of the census field staff and data supervisors who collect and process this information and the Community Liaison Group who run the community engagement programmes. We would particularly like to acknowledge the important contributions of Victoria Nyaga and John Ojal (former HDSS statisticians). This article is published with the permission of the Director of the Kenya Medical Research Institute.

References

1. Africa ECf: **Report on the Status of civil registration and vital statistics in Africa.** ECA, Addis Ababa, 2017.
Reference Source

2. Sankoh O, Dickson KE, Faniran S, *et al.*: **Births and deaths must be registered in Africa.** *Lancet Glob Health.* 2020; **8**(1): e33–e34.
PubMed Abstract | Publisher Full Text

3. Scott JA, Bauni E, Moisi JC, *et al.*: **Profile: The Kilifi Health and Demographic Surveillance System (KHDSS).** *Int J Epidemiol.* 2012; **41**(3): 650–7.
PubMed Abstract | Publisher Full Text | Free Full Text

4. Njuguna P, Maitland K, Nyaguara A, *et al.*: **Observational study: 27 years of severe malaria surveillance in Kilifi, Kenya.** *BMC Med.* 2019; **17**(1): 124.
PubMed Abstract | Publisher Full Text | Free Full Text

5. Hammitt LL, Etyang AO, Morpeth SC, *et al.*: **Effect of ten-valent pneumococcal conjugate vaccine on invasive pneumococcal disease and nasopharyngeal carriage in Kenya: a longitudinal surveillance study.** *Lancet.* 2019; **393**(10186): 2146–2154.
PubMed Abstract | Publisher Full Text | Free Full Text

6. Nokes DJ, Okiro EA, Ngama M, *et al.*: **Respiratory syncytial virus epidemiology in a birth cohort from Kilifi District, Kenya: Infection during the first year of life.** *J Infect Dis.* 2004; **190**(10): 1828–1832.
PubMed Abstract | Publisher Full Text

7. Nokes DJ, Abwao J, Pamba A, *et al.*: **Incidence and clinical characteristics of group a rotavirus infections among children admitted to hospital in Kilifi, Kenya.** *PLoS Med.* 2008; **5**(7): e153.
PubMed Abstract | Publisher Full Text | Free Full Text

8. Berkley J, Mwangi I, Griffiths K, *et al.*: **Assessment of Severe Malnutrition Among Hospitalized Children in Rural Kenya: comparison of weight for height and mid upper arm circumference.** *JAMA.* 2005; **294**(5): 591–7.
PubMed Abstract | Publisher Full Text

9. Wambui KM, Musenge E: **A space-time analysis of recurrent malnutrition-related hospitalisations in Kilifi, Kenya for children under-5 years.** *BMC Nutr.* 2019; **5**: 32.
PubMed Abstract | Publisher Full Text | Free Full Text

10. Uyoga S, Macharia AW, Mochamah G, *et al.*: **The epidemiology of sickle cell disease in children recruited in infancy in Kilifi, Kenya: a prospective cohort study.** *Lancet Glob Health.* The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license. 2019; **7**(10): e1458–e1466.
PubMed Abstract | Publisher Full Text | Free Full Text

11. Kind CJ, Newton CRJC, Kariuki SM, *et al.*: **Prevalence, risk factors, and neurobehavioral comorbidities of epilepsy in Kenyan children.** *Epilepsia Open.* 2017; **2**(4): 388–399.
PubMed Abstract | Publisher Full Text | Free Full Text

12. Kariuki SM, Abubakar A, Kombe M, *et al.*: **Burden, risk factors, and comorbidities of behavioural and emotional problems in Kenyan children: a population-based study.** *Lancet Psychiatry.* 2017; **4**(2): 136–145.
PubMed Abstract | Publisher Full Text | Free Full Text

13. Etyang AO, Munge K, Bunyasi EW, *et al.*: **Burden of disease in adults admitted to hospital in a rural region of coastal Kenya: an analysis of data from linked clinical and demographic surveillance systems.** *Lancet Glob Health.* 2014; **2**(4): e216–24.
PubMed Abstract | Publisher Full Text | Free Full Text

14. Etyang AO, Scott JAG: **Medical causes of admissions to hospital among adults in Africa: A systematic review.** *Glob Health Action.* 2013; **6**: 1–14.
PubMed Abstract | Publisher Full Text | Free Full Text

15. Silaba M, Ooko M, Bottomley C, *et al.*: **Effect of 10-valent pneumococcal conjugate vaccine on the incidence of radiologically-confirmed pneumonia and clinically-defined pneumonia in Kenyan children: an interrupted time-series analysis.** *Lancet Glob Health.* 2019; **7**(3): e337–e346.
PubMed Abstract | Publisher Full Text | Free Full Text

16. Otieno GP, Bottomley C, Khagayi S, *et al.*: **Impact of the Introduction of Rotavirus Vaccine on Hospital Admissions for Diarrhea Among Children in Kenya: A Controlled Interrupted Time-Series Analysis.** *Clin Infect Dis.* 2020; **70**(11): 2306–2313.
PubMed Abstract | Publisher Full Text | Free Full Text

17. Kamau A, Mwangangi JM, Rono MK, *et al.*: **Variation in the effectiveness of insecticide treated nets against malaria and outdoor biting by vectors in Kilifi, Kenya [version 4; peer review: 1 approved, 3 approved with reservations].** *Wellcome Open Res.* 2018; **2**: 22.
PubMed Abstract | Publisher Full Text | Free Full Text

18. Hammitt LL, Crane RJ, Karani A, *et al.*: **Effect of *Haemophilus influenzae* type b vaccination without a booster dose on invasive *H influenzae* type b disease, nasopharyngeal carriage, and population immunity in Kilifi, Kenya: a 15-year regional surveillance study.** *Lancet Glob Health.* 2016; **4**(3): e185–94.
PubMed Abstract | Publisher Full Text | Free Full Text

19. Kamau A, Nyaga V, Bauni E, *et al.*: **Trends in bednet ownership and usage, and the effect of bednets on malaria hospitalization in the Kilifi Health and Demographic Surveillance System (KHDSS): 2008-2015.** *BMC Infect Dis.* 2017; **17**(1): 720.
PubMed Abstract | Publisher Full Text | Free Full Text

20. Otiende M, Bauni E, Nyaguara A, *et al.*: **Replication Data for: Mortality in rural coastal Kenya measured using the Kilifi Health and Demographic Surveillance System: a 16-year descriptive analysis.** Havard Dataverse, 2021.
http://www.doi.org/10.7910/DVN/DDX7OK

21. Ndila C, Bauni E, Mochamah G, *et al.*: **Causes of death among persons of all ages within the Kilifi Health and Demographic Surveillance System, Kenya, determined from verbal autopsies interpreted using the InterVA-4 model.** *Glob Health Action.* 2014; **7**: 25593.
PubMed Abstract | Publisher Full Text | Free Full Text

22. Randall S: **The quality of demographic data on older Africans.** *Demogr Res.* 2016; **34**: 143–174.
Publisher Full Text

23. Tolley HD, Barnes JM, Freeman MD: **Survival Analysis.** *Forensic Epidemiology: Principles and Practice.* Elsevier Inc.; 2016; 261–284.
Publisher Full Text

24. Bocquier P, Ginsburg C, Herbst K, *et al.*: **A training manual for event history data management using Health and Demographic Surveillance System data.** *BMC Res Notes.* 2017; **10**(1): 224.
PubMed Abstract | Publisher Full Text | Free Full Text

25. Bocquier P, Ginsburg C, Collinson MA: **A training manual for event history analysis using longitudinal data.** *BMC Res Notes.* BioMed Central; 2019; **12**(1): 506.
PubMed Abstract | Publisher Full Text | Free Full Text

26. Brewer CA, Pickle L: **Evaluation of methods for classifying epidemiological data on choropleth maps in series.** *Ann Assoc Am Geogr.* 2002; **92**(4): 662–681.
Publisher Full Text

27. Team RC: **R: A Language and Environment for Statistical Computing.** 2021.
Reference Source

28. Masquelier B, Reniers G, Pison G: **Divergences in trends in child and adult mortality in sub-Saharan Africa: survey evidence on the survival of children and siblings.** *Popul Stud (Camb).* 2014; **68**(2): 161–77.
PubMed Abstract | Publisher Full Text

29. Obermeyer Z, Rajaratnam JK, Park CH, *et al.*: **Measuring adult mortality using sibling survival: a new analytical method and new results for 44 countries, 1974-2006.** *PLoS Med.* 2010; **7**(4): e1000260.
PubMed Abstract | Publisher Full Text | Free Full Text

30. Kaufman JS, Asuzu MC, Rotimi CN, *et al.*: **The absence of adult mortality data for sub-Saharan Africa: a practical solution.** *Bull World Health Organ.* 1997; **75**(5): 389–95.
PubMed Abstract | Free Full Text

31. Kharsany AB, Karim QA: **HIV Infection and AIDS in Sub-Saharan Africa: Current Status, Challenges and Opportunities.** *Open AIDS J.* 2016; **10**: 34–48.
PubMed Abstract | Publisher Full Text | Free Full Text

32. Gargano JW, Laserson K, Muttai H, *et al.*: **The adult population impact of HIV care and antiretroviral therapy in a resource poor setting, 2003-2008.** *AIDS.* 2012; **26**(12): 1545–54.
PubMed Abstract | Publisher Full Text

33. Reniers G, Slaymaker E, Nakiyingi-Miiro J, *et al.*: **Mortality trends in the era of antiretroviral therapy: evidence from the Network for Analysing Longitudinal Population based HIV/AIDS data on Africa (ALPHA).** *AIDS.* 2014; **28 Suppl 4**(4): S533–42.
PubMed Abstract | Publisher Full Text | Free Full Text

34. Macharia PM, Giorgi E, Thuranira PN, *et al.*: **Sub national variation and inequalities in under-five mortality in Kenya since 1965.** *BMC Public Health.* 2019; **19**(1): 146.
PubMed Abstract | Publisher Full Text | Free Full Text

35. Quentin W, Abosede O, Aka J, *et al.*: **Inequalities in child mortality in ten major African cities.** *BMC Med.* 2014; **12**: 95.
PubMed Abstract | Publisher Full Text | Free Full Text

36. Kazembe L, Clarke A, Kandala NB: **Childhood mortality in sub-Saharan Africa: cross-sectional insight into small-scale geographical inequalities from Census data.** *BMJ Open.* 2012; **2**(5): e001421.
PubMed Abstract | Publisher Full Text | Free Full Text

37. 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**(12): e936–e945.
PubMed Abstract | Publisher Full Text

38. Joseph NK, Macharia PM, Ouma PO, *et al.*: **Spatial access inequities and childhood immunisation uptake in Kenya.** *BMC Public Health.* 2020; **20**(1): 1407.
PubMed Abstract | Publisher Full Text | Free Full Text

39. O'Meara WP, Bejon P, Mwangi TW, *et al.*: **Effect of a fall in malaria transmission on morbidity and mortality in Kilifi, Kenya.** *Lancet.* 2008; **372**(9649): 1555–1562.
PubMed Abstract | Publisher Full Text | Free Full Text

40. Kleinschmidt I, Schwabe C, Benavente L, *et al.*: **Marked increase in child survival after four years of intensive malaria control.** *Am J Trop Med Hyg.* 2009; **80**(6): 882–8.
PubMed Abstract | Publisher Full Text | Free Full Text

41. Snow RW, Korenromp EL, Gouws E: **Pediatric mortality in Africa: *plasmodium falciparum* malaria as a cause or risk?** *Am J Trop Med Hyg.* 2004; **71**(2 Suppl): 16–24.
PubMed Abstract | Publisher Full Text

42. Rowe AK, Rowe SY, Snow RW, *et al.*: **The burden of malaria mortality among**

African children in the year 2000. *Int J Epidemiol.* 2006; **35**(3): 691–704.
PubMed Abstract | Publisher Full Text | Free Full Text

43. Ross A, Maire N, Molineaux L, *et al.*: **An epidemiologic model of severe morbidity and mortality caused by *Plasmodium falciparum*.** *Am J Trop Med Hyg.* 2006; **75**(2 Suppl): 63–73.
PubMed Abstract | Publisher Full Text

44. Scott JA, Berkley JA, Mwangi I, *et al.*: **Relation between falciparum malaria and bacteraemia in Kenyan children: a population-based, case-control study and a longitudinal study.** *Lancet.* 2011; **378**(9799): 1316–23.
PubMed Abstract | Publisher Full Text | Free Full Text

45. Mogeni P, Williams TN, Fegan G, *et al.*: **Age, Spatial, and Temporal Variations in Hospital Admissions with Malaria in Kilifi County, Kenya: A 25-Year Longitudinal Observational Study.** *PLoS Med.* 2016; **13**(6): e1002047.
PubMed Abstract | Publisher Full Text | Free Full Text

46. Macharia PM, Giorgi E, Noor AM, *et al.*: **Spatio-temporal analysis of *Plasmodium falciparum* prevalence to understand the past and chart the future of malaria control in Kenya.** *Malar J.* 2018; **17**(1): 340.
PubMed Abstract | Publisher Full Text | Free Full Text

47. Cowgill KD, Ndiritu M, Nyiro J, *et al.*: **Effectiveness of *Haemophilus influenzae* type b Conjugate vaccine introduction into routine childhood immunization in Kenya.** *JAMA.* 2006; **296**(6): 671–678.
PubMed Abstract | Publisher Full Text | Free Full Text

Open Peer Review

Current Peer Review Status: ✔ ? ✔ ✔

Version 2

Reviewer Report 04 July 2023

https://doi.org/10.21956/wellcomeopenres.20971.r60389

✔ **Alain Vandormael**
Heidelberg University, Im Neuenheimer Feld, Heidelberg, Germany

Introduction

1. The introduction is clear and well-written.

Methods

1. The authors write “Subsequently, the population has been continuously monitored for births, pregnancies, deaths, and migration events in re-enumeration rounds occurring approximately every 4 months, and the mapping was updated in 2017.” Is this first sentence saying that all data points (births, pregnancies, deaths) are collected from the re-enumeration rounds, or just the migration data?

2. In the next paragraph, the authors write that there are 4 points of contact, where re-enumeration rounds refer to household visits. Births and pregnancies information is collected from the maternity wards. This sentence suggests that not all data is collected from the re-enumeration rounds (see comment #1). Please fix this to make the sentence and data collection methods clearer.

3. What are the four-month re-enumeration rounds referring to specifically? The average time between household visits or the average time between data points (not just household data from a key informant but also from clinics and maternity wards). This is not clear.

4. Do field workers visit only the households every for months, or do they also go to the wards, clinics and vaccination sites every four months. This is not clear.

5. The authors note that all household members can be absent during a household visit. No data or information is given on participation rates in the surveillance program. Can more be said about this? Is the household participation rate low, moderate or good? Is there a paper from the Kilifi surveillance that documents and quantifies missed household visits and non-response rates in more detail? For example, at the Welcome Trust Surveillance site in northern Kwazulu-Natal, South Africa, there is the paper by Larmarange *et al.*[1] that does

this.

(I am not asking for such a study to be done.) Has a previous study been done that you can reference? The reader will observe the high rates of missed HIV test dates and household visits in South Africa Welcome Trust site. Is it similar in current surveillance site? Can the authors say something about missing data in the Methods section and how it was dealt with.

Statistical Methods

1. Overall, the authors provide a clear description and justification of the statistical methods used.

2. How was missing data dealt with in the statistical methods? The authors talk about dealing with outmigration, but not about missed household visits, etc. I expect there must be a non-negligible amount of missing data.

3. How was data from the clinics and maternity wards matched, merged, and collated with data from the household visits. What was the match rate? For example, how was participant data from medical records linked to household data records. What % of these records, identifiers could be matched? Can more be said about this data quality aspect of the study.

Results

1. The results are clearly presented. The plots showing trends and the maps showing changes in outcomes over time are informative.

2. Minor comments: Median survival increased from 65 *years" in 2003 to 77 *years* in 2018 (pg 5).

Discussion

1. The authors write: "This pattern is consistent with mortality sex ratios before and after the rapid initiation of HIV care and antiretroviral therapy in LMICs between 2003 and 2018 with evidence showing that women have benefitted more from the expansion of HIV treatment programs than men." It is not possible that the references given (32,33) can be used to support this claim, since they were published well before the 2003-2018 period end. Moreover, in a generalized epidemic with heterosexual transmission, we would expect men to benefit more if HIV treatment is increased among women. Indeed, recent evidence[2] suggests that men have been benefitting more from the increased HIV care expansion among women from 2013 to 2018.

2. The authors write: "In Kilifi, the prevalence of HIV infection among women attending antenatal at KCH ranged between 3.8–4.4% between 2005–2009 and 2.0–3.7% between 2010–2016 with a clear decline from 20105." Can the authors elaborate more on this HIV prevalence trend. If women are living longer over the study period, and HIV care is being expanded among women, then we would expect HIV prevalence to increase (not decrease) since HIV-positive women on treatment are expected to live longer. For the relationship between ART coverage and HIV prevalence, see for example: Zaidi *et al.,*[3].

3. This paper was written in 2023. Are there studies that have been undertaken in the current surveillance area that report on ART coverage and other public health services such as voluntary medical male circumcision (other than malaria prevention scale-up). While I understand that this is a descriptive analysis or mortality trends in the surveillance area, it

would be useful for the reader if the authors can hypothesize why they have seen improvements in mortality over the last 16 years. It is highly likely that the scale-up of HIV treatment and prevention strategies is an important part of this story. The authors touch on this point lightly and with respect to LMICs, but what about saying more about the potential impact of these programs specifically in the surveillance area.

References

1. Larmarange J, Mossong J, Bärnighausen T, Newell ML: Participation dynamics in population-based longitudinal HIV surveillance in rural South Africa.*PLoS One*. 2015; **10** (4): e0123345 PubMed Abstract | Publisher Full Text
2. Vandormael A, Akullian A, Siedner M, de Oliveira T, et al.: Declines in HIV incidence among men and women in a South African population-based cohort.*Nat Commun*. 2019; **10** (1): 5482 PubMed Abstract | Publisher Full Text
3. Zaidi J, Grapsa E, Tanser F, Newell ML, et al.: Dramatic increase in HIV prevalence after scale-up of antiretroviral treatment.*AIDS*. 2013; **27** (14): 2301-5 PubMed Abstract | Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
No

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: HIV epidemiology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 04 July 2023

https://doi.org/10.21956/wellcomeopenres.20971.r60394

Ian Cook
University of Limpopo, Mankweng, Limpopo, South Africa

General comments:

Well written and constructed paper - achieves a balance between detail and overview. Also, caters for a specialised reader as well as a more general readership. Notable, the number of publications which have emanated from the KDHSS. A large database with almost 300 000 residents in 2018, 700 000 individuals and nearly 4 million p-y of observation, analysed over 16 years.

Specific comments:

- It might be useful (if even in an appendix) to provide a map(s) (country and county), identifying the site and important landmarks (hospitals/towns, cities/major transport routes etc) in the site area. And the location of the other Kenyan INDEPTH sites.

- The authors mention the use of INDEPTH methodology, but do not explicitly state that the KHDSS is part of the INDEPTH network of sites (http://www.indepth-network.org/member-centres). I note there are five HDSS sites in Kenya (KHDSS included) - yet the discussion does not explicitly compare (even briefly) the results with that of other Kenyan INDEPTH sites. If there are no comparable Kenyan INDEPTH data, the authors might consider explicitly stating this.

- Sufficient figures, which are clear and easily interpretable, to highlight trends.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
I cannot comment. A qualified statistician is required.

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Public Health, Physical Behaviours Epidemiology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 14 January 2022

https://doi.org/10.21956/wellcomeopenres.19134.r47348

?

Momodou Jasseh
Medical Research Council Unit The Gambia at LSHTM, Fajara, The Gambia

Mortality in rural coastal Kenya measured using the Kilifi Health and Demographic Surveillance System: a 16-year descriptive analysis

The authors claim the paper to be the first attempt ever to systematically characterise general mortality in the Kilifi HDSS since it was established in 2000. Being an area that has hosted a significant number of important trials and health interventions of the KEMRI-Wellcome Trust Research Programme, and with access to tertiary health care through the Kilifi County Hospital, a systematic description of levels and trends in all-cause mortality in the area is necessary and relevant; and the attempt is therefore justified.

Whilst the paper is generally well-written, it can be enhanced further both in terms of content and presentation as indicated in the following queries.

GENERAL QUERIES

1. The frequent reference made to supplementary tables and figures in the paper makes reading it a rather arduous task. If tables and figures are important enough to be presented as part of the results, then they should appear in the paper accordingly. The authors should therefore decide what information to include in the paper and how they should be lucidly presented to enhance clarity and ease comprehension.

2. None of the mortality rates estimated and shown in the tables and graphs are presented with Confidence Intervals (95% CI). It is recommended that these be provided for all indicator estimates presented in tables, and as CI bars in levels/trends graphs. Not only will they add value to the results of the study but will also enable the reader to make inferences/deductions from the presented results beyond what the authors have reported, especially – for example – with respect to confirming statistically significant differences

between indicators for different time periods.

3. I observe the interchangeable use of the terms “rates” and “ratios”. Since mortality indicators are conventionally estimated as rates, I will suggest that the term “rates” be used consistently in the paper.

4. Too many methods are used to characterise mortality for the first time in the Kilifi HDSS area. The purpose of the paper is mortality estimation and validation of the KHDSS data, and not comparison of methods using the same dataset. Therefore, it may be worthwhile to adopt the conventional methods of mortality estimation to effectively describe the characteristics of the region’s mortality experience for the stated period; as well as facilitate easy comparison with estimates from other data sources, and for other regions in Kenya. For instance, comparisons of life expectancy estimates from the Life Table method with similar estimates from census or DHS sources provides more analytical value in assessing the quality of the KHDSS data than comparing them with Kaplan-Meier-derived mean life expectancies from the same dataset. The Life Table method also provides the opportunity to cross-check and validate derived age-specific mortality rates.

SPECIFIC QUERIES

1. *Page 4, left column, fifth para*: Remove the word “survival” from the sub-title. It can be used interchangeably with “mortality” and the choropleth maps refer to mortality rates over time.

2. *Page 4, left column, fifth para*: Provide a justification for applying the direct standardization of the populations against the 2011 KHDSS age structure. If the idea is to use the mid-period age structure for this purpose, then that of 1st January 2013 should be used instead.

3. *Pages 5 & 6*: Consider merging Tables 1 and 2 to provide comprehensive information on births, deaths, PYO and corresponding mortality rates (with 95% CI) by age-group and period. The childhood mortality estimates (i.e. neonatal, post-neonatal, infant [<1 yr]), child [1-4 yrs] and under-5) can be expressed in both per 1,000 PYO and per 1,000 live births where applicable. The authors are urged to reconsider the method used in estimating the childhood mortality rates presented in Table S2. These rates should be equivalent to the probabilities of dying within the respective age brackets, and can be estimated more accurately using the Lexis Diagram method, for instance. The under-5 mortality rate should not be the sum of the infant mortality (<1) and child mortality (1-4), but rather the complement of the product of the probability of surviving the first year of life and the probability of surviving to exact age 5 thereafter. The authors should therefore review their method of analysis to provide accurate measures of childhood mortality indicators. In fact, the mortality rate for children aged 1-4 years cannot be expressed in “per 1,000 live births”, because it refers to a population defined by having already survived the first year of life.

4. *Page 5, left column, last para*: The statement that a seasonal pattern appears only in children aged 1-4 years in the period 2003-2006 is speculative. Whether it is real or a data artefact can be ascertained by adopting more robust statistical analysis. An inclusion of 95% CI bars in Figure 3 would have enabled the reader to make an inference to confirm or dismiss such a seasonal mortality pattern claim.

5. *Page 10, right column, second para*: The explanation linking oestrogen levels and female life expectancy is beyond the scope of the paper and should be removed.

6. Whilst sufficient comparison is done between KHDSS childhood mortality indicators with estimates from other sources and regions, not much similar comparison is attempted for adult mortality.

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Demography

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 20 Jan 2023

Mark Otiende

Response to the reviewer
We thank the reviewer for his interest in our paper and for taking the time to review it thoroughly. We have given point-by-point responses in italics.

Mortality in rural coastal Kenya measured using the Kilifi Health and Demographic Surveillance System: a 16-year descriptive analysis

The authors claim the paper to be the first attempt ever to systematically characterise general mortality in the Kilifi HDSS since it was established in 2000. Being an area that has hosted a significant number of important trials and health interventions of the KEMRI-Wellcome Trust Research Programme, and with access to tertiary health care through the Kilifi County Hospital, a systematic description of levels and trends in all-cause mortality in

the area is necessary and relevant; and the attempt is therefore justified.

Whilst the paper is generally well-written, it can be enhanced further both in terms of content and presentation as indicated in the following queries.

GENERAL QUERIES

1. The frequent reference made to supplementary tables and figures in the paper makes reading it a rather arduous task. If tables and figures are important enough to be presented as part of the results, then they should appear in the paper accordingly. The authors should therefore decide what information to include in the paper and how they should be lucidly presented to enhance clarity and ease comprehension.

   As mentioned in the review, Kilifi HDSS is the setting for a large number of studies of observational epidemiology and public health interventions and scientists reading these studies may wish to interrogate the background mortality experience of the population from a variety of different perspectives. This is what has motivated our analyses – this is expressed in the introduction. We divided the results between the main paper and a supplement to produce a main paper that was accessible to the general reader and simultaneously provide a compendium of tabulation sufficiently comprehensive to satisfy the more detailed analyst. We have decided what information we wish to emphasise and provided signposts to the supplementary data merely to inform the more curious reader.

2. None of the mortality rates estimated and shown in the tables and graphs are presented with Confidence Intervals (95% CI). It is recommended that these be provided for all indicator estimates presented in tables, and as CI bars in levels/trends graphs. Not only will they add value to the results of the study but will also enable the reader to make inferences/deductions from the presented results beyond what the authors have reported, especially – for example – with respect to confirming statistically significant differences between indicators for different time periods.

   In our analysis, the Kilifi HDSS is both the population of interest and the sample population. Since we are not inferring to a larger population, confidence intervals for our mortality rates estimates are unnecessary. However, we agree with the reviewer that it is important to rule out month-to-month fluctuations (e.g. due to conceptions, temperature, rainfall, crop yields, road access to hospital, vaccination coverage, etc) when comparing between time periods. To rule out random fluctuations (Table 2 and Figure 3a), we have fitted a negative binomial model with yearly rates as the outcome (yearly rates eliminate the issue of seasonality) and period as the explanatory variable. For all age groups under 55 years, we observed significantly lower mortality rates in the latter periods compared to the first period (2003-2006). We have added a supplementary table (Table S2) showing these comparisons for all age groups.

3. I observe the interchangeable use of the terms “rates” and “ratios”. Since mortality indicators are conventionally estimated as rates, I will suggest that the term “rates” be used consistently in the paper.

We have emphasised the calculation of rates because we have a sophisticated dataset with events linked to person years at risk. Much demographic data does not have this linkage and many of the estimates of mortality are calculated as the ratio of the number of deaths in any one age group in one place in one year to the number of live births in the same place in the same year. These ratios are used as an approximation of the cumulative mortality incidence but they have a number of shortcomings including the fact that where the birth rate fluctuates annually, they misrepresent the denominator for any group beyond infancy, and the fact that they calculate the mortality based on the starting population not the person years at risk. From this perspective they are risks (not rates) but the lack of correspondence between the numerator and denominator negates this description and so we, and others[1,2], have referred to these as ratios. As much demographic data is captured with this method we have calculated the ratios in Kilifi for wider comparison. We have not used the terms interchangeably and in fact we defined their separate and distinct use in the statistical methods (Age-sex mortality profile); the first paragraph defines the calculation of rates, the second the calculation of (and justification for) ratios.

4. Too many methods are used to characterise mortality for the first time in the Kilifi HDSS area. The purpose of the paper is mortality estimation and validation of the KHDSS data, and not comparison of methods using the same dataset. Therefore, it may be worthwhile to adopt the conventional methods of mortality estimation to effectively describe the characteristics of the region's mortality experience for the stated period; as well as facilitate easy comparison with estimates from other data sources, and for other regions in Kenya. For instance, comparisons of life expectancy estimates from the Life Table method with similar estimates from census or DHS sources provides more analytical value in assessing the quality of the KHDSS data than comparing them with Kaplan-Meier-derived mean life expectancies from the same dataset. The Life Table method also provides the opportunity to cross-check and validate derived age-specific mortality rates.

   In writing this paper our purpose was two-fold. We wanted to provide a digestible description of the mortality patterns and trends for those who have read the many epidemiological studies and intervention trials that have taken place in Kilifi and we wanted to provide the comprehensive demographic tables to allow demographers to compare our data with other published estimates. As many of the prior studies in Kilifi have been epidemiological, the readers are familiar with survival analyses and the visual efficiency with which a Kaplan Meier curve can convey information on (instantaneous) mortality rates by age. Conventional analyses using life tables, by contrast, provide tabulated discrete data in many strata that are less easy for the general reader to assimilate rapidly but are necessary for the analysis of standard demographic outputs. In line with our purpose to make the paper digestible, we have emphasised the presentation of survival curves in the paper and in line with our purpose to be comprehensive we have included a full set of life tables in the appendix. We accept that, to one audience segment, the survival analyses may seem superfluous, and to the other the life tables may seem inaccessible but we aim to reach both and we invite our readers to take note, that the two approaches provide slightly different estimates of life expectancy – and we discuss this briefly.

SPECIFIC QUERIES

1. *Page 4, left column, fifth para*: Remove the word "survival" from the sub-title. It can be used interchangeably with "mortality" and the choropleth maps refer to mortality rates over time.

   Following directly from the response above – we have tried to signpost which analytic approach is being used at each point. The sub-header 'Survival' precedes a paragraph explaining the survival analysis of the mortality data and a reference to six Kaplan Meier survival curves. Later, under the heading Geographic heterogeneity in survival and mortality we signpost the variation in survival (from the survival analyses which we detail in the supplement) and then present here the location-specific mortality rates in the choropleth maps. Our judgment is that, whilst survival curves describe mortality by age efficiently, mortality rates provide a more efficient metric to understand the variation of mortality cartographically.

2. *Page 4, left column, fifth para:* Provide a justification for applying the direct standardization of the populations against the 2011 KHDSS age structure. If the idea is to use the mid-period age structure for this purpose, then that of 1st January 2013 should be used instead.

   We stand by our calculation – for a period beginning 1st January 2003 and ending 31st December 2018, a mid-point estimate of 1st January 2011 is perfectly reasonable (it could be argued that 31st December 2010 may be more accurate – in days…) The purpose is to find a reasonable reference population for a standardisation exercise – it is not necessary to fix it to the mid-study point, though this is what we did.

3. *Pages 5 & 6:* Consider merging Tables 1 and 2 to provide comprehensive information on births, deaths, PYO and corresponding mortality rates (with 95% CI) by age-group and period. The childhood mortality estimates (i.e. neonatal, post-neonatal, infant [<1 yr]), child [1-4 yrs] and under-5) can be expressed in both per 1,000 PYO and per 1,000 live births where applicable. The authors are urged to reconsider the method used in estimating the childhood mortality rates presented in Table S2. These rates should be equivalent to the probabilities of dying within the respective age brackets, and can be estimated more accurately using the Lexis Diagram method, for instance. The under-5 mortality rate should not be the sum of the infant mortality (<1) and child mortality (1-4), but rather the complement of the product of the probability of surviving the first year of life and the probability of surviving to exact age 5 thereafter. The authors should therefore review their method of analysis to provide accurate measures of childhood mortality indicators. In fact, the mortality rate for children aged 1-4 years cannot be expressed in "per 1,000 live births", because it refers to a population defined by having already survived the first year of life.

   We have extended Table 2 to include age and period specific numbers of deaths and pyo in addition to rates. In our selection of results we have emphasised mortality rates (over ratios) for the reasons outlined in general query #3 above and have only presented the infant mortality ratio, under 5 mortality ratio etc for the sake of comparison with other demographic sources, and only in the supplement. We agree with the reviewer that these

ratios provide a metric which has no basis in reality ('the mortality rate in children aged 1-4 years cannot be expressed...') but this analytic approach is, nonetheless, in widespread use and is frequently taught (see for example the definition of 'Post-neonatal mortality rate' at https://www.cdc.gov/csels/dsepd/ss1978/lesson3/section3.html). For exact probabilities of dying from one age point to the next, we refer our readers to the Life Tables in the supplement.

4. *Page 5, left column, last para*: The statement that a seasonal pattern appears only in children aged 1-4 years in the period 2003-2006 is speculative. Whether it is real or a data artefact can be ascertained by adopting more robust statistical analysis. An inclusion of 95% CI bars in Figure 3 would have enabled the reader to make an inference to confirm or dismiss such a seasonal mortality pattern claim.

   In addition to the visual interpretation of the data (a seasonal pattern appearing only in in children aged 1-4 years), results from a negative binomial regression model confirm a significant interaction between season and period in children aged 1-4 years only. We have reported the same in the in the methods, results and supplement.

5. *Page 10, right column, second para:* The explanation linking oestrogen levels and female life expectancy is beyond the scope of the paper and should be removed.

   We have removed it.

6. Whilst sufficient comparison is done between KHDSS childhood mortality indicators with estimates from other sources and regions, not much similar comparison is attempted for adult mortality.

   There is a lot less data on adult mortality compared to child mortality which has made it difficult to interpret these estimates. We have made this point in the discussion (paragraph 2) and also cited adult mortality estimates derived from DHS sibling survival data.

Competing Interests: No competing interest

Reviewer Report 16 December 2021

https://doi.org/10.21956/wellcomeopenres.19134.r47346

Stephen Gordon
[1] Department of Clinical Sciences, Liverpool School of Tropical Medicine, Liverpool, UK

[2] Malawi Liverpool Wellcome Programme, Blantyre, Malawi

General comments: Elegant, detailed and clear account of a large well conducted study. Many high quality publications have emerged from the Kilifi DHSS which has become a model for many others.

Specific comments
Title and Abstract are clear. The abstract would be improved by including the size of the dataset.

Introduction: Clear, but could be improved by a little bit of geography to explain the terrain, who the people are, a bit more colour!

Methods: Clear, and a good explanation of how to deal with in- and out-migration.

Results: 22207 deaths and 3.9m pyo is a very big study. It is good to see such a precise data set to support other data based on less solid observations. I note the lack of seasonality, which is different from that published from Gambia. Would you like to comment? Geographical variation is found everywhere (e.g. in UK) and often relates to poverty. Could you look at poverty, clean water or distance to health facilities in your analysis?

Discussion: Nice discussion of HIV, vaccination and gender. I'm curious about the north/south divide that has emerged across the creek with the north showing more improvement. Can you comment? Good comparison with other data sources and explanation of differences.

Figures: elegant and helpful.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
I cannot comment. A qualified statistician is required.

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Respiratory Medicine and general internal medicine, work in Low Income

Countries.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.