

Risk factors associated with rural water supply failure: a 30-year retrospective study of handpumps on the south coast of Kenya

Tim Foster^{1,2,}, Juliet Willetts¹, Mike Lane³, Patrick Thomson², Jacob Katuva² & Rob Hope²*

¹Institute for Sustainable Futures, University of Technology Sydney, 235 Jones St, Ultimo 2007, NSW, Australia

²School of Geography and the Environment and Smith School of Enterprise and the Environment, Oxford University, South Parks Road, Oxford, OX1 3QY, United Kingdom

³Rural Focus Ltd, Box 1011-10400, Nanyuki, Kenya

* Corresponding Author

Tim Foster

Address: Institute for Sustainable Futures, University of Technology Sydney, 235 Jones St, Ultimo 2007, NSW, Australia

Email: tim.foster@uts.edu.au

KEYWORDS: survival analysis; rural water supply; handpumps; groundwater; Kenya

ABSTRACT

An improved understanding of failure risks for water supplies in rural sub-Saharan Africa will be critical to achieving the global goal of safe water for all by 2030. In the absence of longitudinal biophysical and operational data, investigations into water point failure risk factors have to date been limited to cross-sectional research designs. This retrospective cohort study applies survival analysis to identify exogenous factors that predict failure risks for handpumps installed on boreholes along the south coast of Kenya from the 1980s. The analysis is based on a unique data set linking attributes of more than 300 water points at the time of installation with their operational lifespan over the following decades. Cox proportional hazards and accelerated failure time models suggest water points are at higher risk of failure and have shorter lifespans when water supplied is more saline, static water level is deeper, and groundwater is pumped from an unconsolidated sand aquifer. The risk of failure also appears to grow as distance to spare part suppliers increases. To bolster the sustainability of rural water services and ensure no community is left behind, post-construction support mechanisms will need to mitigate heterogeneous environmental and geographical challenges. Further studies are needed to better understand the causal pathways that underlie these risk factors, in order to inform policies and practices that ensure water services are sustained even where unfavourable conditions prevail.

1. INTRODUCTION

Water point sustainability has long been an elusive goal in rural sub-Saharan Africa. Studies and monitoring data have repeatedly revealed a considerable proportion of water points – especially wells and boreholes equipped with handpumps – in a state of disrepair (RWSN, 2009; Jimenez & Perez-Foguet, 2011; Foster, 2013). The human development implications of this situation remain unquantified, but the health and welfare consequences are likely to be substantial. With almost one million handpumps installed across the continent (MacArthur, 2015), it is plausible that tens of millions of rural Africans bear the burden of non-functional systems at any point in time.

A burgeoning body of literature has sought to unravel the predictors and causes of water point operation and maintenance failures. Methodologies and diagnostic frameworks have included multivariable statistical analysis – usually of relatively large water point mapping datasets – to understand determinants of functionality status (Whittington et al. 2009; Foster 2013; Fisher et al. 2015; Cronk & Bartram, 2017), detailed technical assessments of failure modes (Harvey, 2001; Bonsor et al., 2015), in-depth socio-technical root cause analysis (Bonsor et al., 2015), and systems dynamic modelling (Walters & Javernick-Will, 2015). Each approach has its attendant strengths and weaknesses, bearing in mind that they each seek to answer different questions with varying levels of precision.

Multivariable logistic regression has been a commonly employed statistical technique to empirically assess water point functionality risk factors (Foster, 2013; Fisher et al., 2015; Whittington et al., 2009; Cronk & Bartram, 2017). More recently, Bayesian network modelling has been proposed as an approach better able to accommodate the interdependent nature of explanatory variables (Fisher et al., 2015; Cronk & Bartram, 2017). A key shortcoming of these cross-sectional analyses is the imperfect (albeit convenient) binary

functionality status indicator that has been applied as the outcome variable of interest. The inadequacies of a simplistic, point-in-time snapshot of ‘functionality’ are well documented (Thomson et al., 2012; Carter & Ross, 2016), and there is a need for more nuanced examination that distinguishes between those non-functional water points that have long been abandoned and those which are temporarily broken down but likely to be repaired in the near future.

A related drawback of cross-sectional studies utilising a functionality outcome variable is their susceptibility to reverse causation. Take, for example, the collection of revenue from water users, which is a commonly employed explanatory variable and has emerged as a significant determinant of water point functionality in several cross-sectional studies. A lack of user fees is clearly a plausible reason why faulty water points might go unrepaired; however, this association could also be linked to water points initially falling into disrepair for non-financial reasons, which subsequently causes water committees to abandon user fee collection. In other words, the outcome variable might in some cases precede and influence an explanatory variable rather than the other way around.¹ Cross-sectional water point data sets are typically not well suited for assessing the extent to which such factors are a precursor to water point failure or simply a consequence.

A further weakness of functionality studies drawing on large cross-sectional datasets has been their tendency to omit important groundwater characteristics specific to each water point, such as depth and water quality parameters. This is partly due to practical constraints: assessing water quality for a non-functional handpump often requires the pump to be

¹ The interpretation of such associations ultimately depends on the specific point in time to which respondents refer when they provide information forming the basis of explanatory variables, and this in turn may hinge on the way a question is worded by an enumerator. Other explanatory variables that could potential serve as both drivers and consequences of functionality status include a water committee’s composition and level of activity, user group size, and number of water sources in a community.

84 disassembled, as does the measurement of static water level regardless of operational status
85 (with some exceptions)². Collecting such data for handpump water supplies post-installation
86 can therefore be a laborious and expensive process. Although there are alternative measures
87 that have been used as proxies, they tend to be imprecise. For example, Fisher et al. (2015)
88 and Cronk and Bartram (2017) incorporated groundwater storage, depth and yield into their
89 analysis by overlaying spatial datasets bearing a 5km resolution, while Foster (2013) relied
90 on user perceptions of water quality.

91 In contrast, Bonsor et al. (2015) have proposed a toolbox of approaches and a diagnostic
92 framework that enables a more comprehensive assessment of underlying causes of failure,
93 and more precisely considers the role of hydrogeological characteristics. Qualitative
94 narratives are an important component of this process of enquiry in order to untangle the
95 longitudinal and interlinking sequence of events that lead to a water point failure (Carter &
96 Ross, 2016). However, the flipside to these more rigorous investigative processes is that they
97 are likely to be more time consuming and costlier than analysis of water point datasets
98 collected through routine monitoring efforts. The financial implications may also result in
99 smaller sample sizes, making it more difficult to draw definitive conclusions of broad
100 application.

101 This study utilises a research design and analytical approach that avoids some of the
102 abovementioned limitations of statistical analysis of large cross-sectional datasets, but can
103 still be applied where resource or time constraints might prevent a thorough water point-by-
104 water point root cause analysis. We apply a set of statistical techniques collectively known as
105 survival analysis to the context of water supply systems in rural Kenya. Specifically, we
106 employ Kaplan-Meier estimates (Kaplan & Meier, 1958), Cox proportional hazards models

² Some handpump models do include an inspection panel that allows for water level measurements to be taken without the need to remove down-the-hole pump components

(Cox, 1972), and accelerated failure time models in order to identify risk factors associated with the failure of handpump water supplies on the South Coast of Kenya over the course of several decades. The analysis exploits data drawn from water point installation records documented during the 1980 and 1990s, and a follow-up assessment of their location, operational status, and lifespan in 2013.

Survival analysis adopts ‘time until an event occurs’ as its outcome variable of interest. The techniques that fall within the survival analysis umbrella have been used extensively in medical literature due to their ability to accommodate right-censored data – that is, subjects that have not yet undergone the event of interest (such as disease onset, remission or death) by the end of an observation period (see e.g. Clark et al., (2003); Bradburn et al., (2003)). In that respect, the technique can be used to examine water point lifespans, even if a subset of those water points are still operational and their ultimate survival times are not yet known. As well as applying analytical techniques that have not yet been brought to bear on rural water sustainability issues, by matching installation records with subsequent survival times the investigation is able to avoid concerns relating to reverse causation, and consider the influence of hydrogeological conditions irrespective of a water point’s ultimate functionality status.

2. STUDY SITE

The study took place in Kwale County, a predominantly rural region situated on the south coast of Kenya. Kwale has a unique place in the history of rural water supply programming, as it was the site for the first large scale deployment of the Afridev handpump. The Afridev is a lever-action reciprocating handpump originally designed to be maintained at the village-level, and capable of a pumping lift of up to 45m (Baumann & Furey, 2013). It is now a

mainstay water supply technology across rural sub-Saharan Africa, being the favoured handpump model in seven countries, and common in at least 12 others (MacArthur, 2015). A small field trial of the Afridev commenced in the then Kwale District in 1983. The positive results arising from the pilot spurred a district-wide roll-out of the technology from 1985 under the banner of the Kwale Water and Sanitation Project (KWSP) (Narayan-Parker, 1988; McCommon et al., 1990). The Swedish International Development Cooperation Agency (SIDA) played a critical role in financing the programme's expansion. The vast majority of handpumps were fitted on boreholes, and County Government records suggest more than 550 had been installed by the mid-1990s (Figure S1 in Supplementary Material). The handpump-equipped boreholes installed under the KWSP (hereafter termed 'KWSP water points') were distributed across a number of geological formations (Figure S2 in Supplementary Material). These include Pleistocene coral limestones, Pleistocene and Pliocene sands (Kilindini and Margarini formations), Triassic sandstones and shales (Mazeras, Mariakani, and Maji-ya-Chumvi formations), and to a smaller extent, igneous intrusives (Caswell, 1953). In addition to installing hardware, the KWSP also made concerted efforts to establish community-based water committees, and provide training on administrative, financial and technical aspects of operation and maintenance.

For several reasons, Kwale provides an instructive setting in which to assess water point lifespans, and the exogenous predictors thereof. First, as one of the earliest programmes of its kind in sub-Saharan Africa, it allows for an analytical frame spanning three decades. Second, as a large number of water points were installed under the same programme, some implementation fundamentals are held relatively constant – such as technology and approach to community participation – as compared with many other rural water landscapes that are

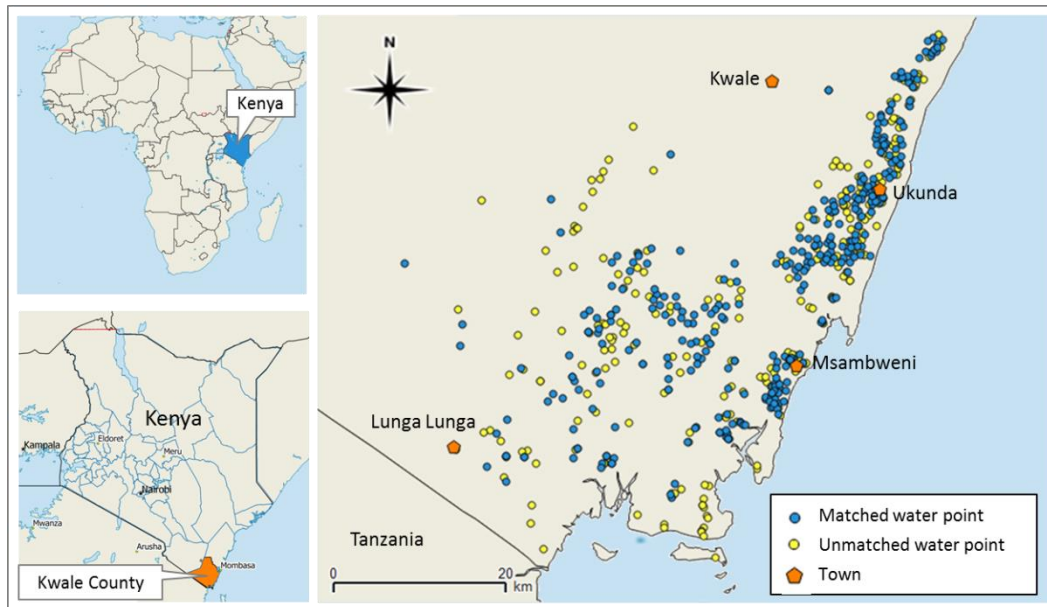
more fragmented and less harmonized. The communities in Kwale therefore shared the same technical and institutional starting point for operation and maintenance (e.g. Afridev handpumps, local mechanics, trained water committees), allowing for a clearer distillation of the effect of exogenous factors. Crucially – and rather uniquely – this coordinated approach also resulted in a centralized repository of installation records containing key characteristics of the water points installed.

3. METHODS

3.1 Data collection

A water point mapping exercise was carried out in Kwale in 2013 with the aim of locating Afridev handpump installations in Msambweni, Matuga and Lunga Lunga sub-counties. This data collection took place as part of a wider research programme into rural water sustainability (see Foster & Hope, 2016; 2017). During this process, a range of geographical information on water points was captured, including community name, GPS coordinates, and – where extant – unique identification codes and installation dates inscribed on auxiliary concrete works such as aprons and washstands. In total, 571 Afridev installation sites were located (Figure 1), of which 314 (55.0%) were functional, 238 (41.7%) were non-functional (i.e. not producing water at the time of inspection), and 19 (3.3%) had been replaced by a motorized pump. For those water points that were deemed non-functional, the respondent was also asked to estimate the duration of the current breakdown. To mitigate recall bias, the same question was also asked of other surrounding households (mean 4.8 households per water point), and the median value of the breakdown duration was determined.

Figure 1. Study site and distribution of Afridev handpumps in Kwale County, 2013



Subsequent to the water point mapping, original installation records kept by the Kwale County government were consulted. These written records contained details for 580 handpump-equipped boreholes installed between 1983 and 1995. This included information on community name, water point identification code, drilling commencement and completion dates, borehole depth, static water level, yield, and electrical conductivity. Twenty-one water points included in the records were located outside of the study area, and were excluded from the study. Characteristics of the remaining 559 water points at the time of installation are presented in Table 1.

Table 1. Mean characteristics of water points equipped with Afridev handpumps included in Kwale County Government records (1983-1995) (n=559)

	Mean	SD	Median	Min	Max
Borehole depth (m)	36.2	18.7	32.0	7.0	102
Static water level (m)	17.0	9.3	16.8	0.9	60.4
Electrical conductivity (µS/cm)	1097	1113	690	60	6100
Yield (l/s)	0.73	0.52	0.67	0.08	5.5
Year of installation	1989	3.1	1989	1983	1995

Note: All characteristics relate to the time of installation

A process was then carried out to match water points identified in the 2013 inventory with those contained in the installation records. Water points were matched based on either the identification code or a unique location name. In total, 337 water points appearing in the installation records were able to be linked with specific water points located during the 2013 mapping process (Table S1 in Supplementary Information). These matched water points constituted 60% of the 559 water points recorded in the installation records, and 59% of the 571 Afridev installation sites located during the water point mapping in 2013. Based on the distinguishing physical features of KWSP handpump installations (i.e. conical concrete pedestal, concrete wash basins), we estimated that a further 113 water points located in the 2013 inventory were KWSP water points; however, these water points could not be traced to specific water points in the installation records. This left 109 water points from the installation records that were not located at all in 2013. It is plausible that some of these

residual water points had little or no extant above-ground infrastructure, and hence have been non-functional for years or decades. Some may have also been upgraded or replaced with more advanced water supply systems in a way that left no evidence of the original handpump installation.

3.2 Variables

The outcome variable of interest was operational lifespan. This was defined as the number of years between installation and the point at which a water point had been non-functional for a year (hereafter called “failure”). Those water points that were functional or non-functional for less than one year at the time of the water point mapping were therefore right-censored – in other words, because they were operational or recently operational, their lifespan up until the point of failure was not yet known. Breakdown duration of at least one year was chosen as the threshold for failure so it would not inadvertently capture water points left unused and in a state of disrepair on a seasonal basis.

Nine exogenous explanatory variables were included in the analysis (Table S2 in Supplementary Information). The variables were selected based on hypotheses that they would influence a water point’s risk of failure, and they were either documented for a large number of water points in the installation records, or could be deduced from the location of the water point. The variables related to a water point’s hydrogeological characteristics (static water level; yield; geological formation; electrical conductivity), time of installation (rainfall season); and its location (relative to year-round fresh surface water; relative to spare parts suppliers; settlement type; and whether the water point was institutional or communal).

With the exception of geological formation, hydrogeological variables were obtained from the installation records. In order to control for seasonal variation in hydrogeological variables, drilling completion date was converted into a dichotomous seasonal variable with

the ‘dry’ period defined as January to April – the time of year when groundwater levels tend to be at their deepest in Kwale (Thambu, 1987) – with the remainder of the year constituting ‘wet’ season. The type of user group (community vs institution) was determined from the name of the water point entered into the installation records. If the water point name specifically referred to a school, health facility or police station then it was deemed institutional; if not, it was assumed to be a community facility. It is acknowledged that in reality such a distinction might not always be clear cut – community members might use an institution’s water point and vice versa. A variable indicating distance to spare parts was formulated by calculating the straight-line distance between each water point and Ukunda town, the major commercial centre from which spare parts have been supplied. Geological formations were consolidated into three categories, namely ‘corals’ (Pleistocene corals), sands (Pleistocene and Pliocene sands), and ‘other’ (Mazeras sandstones, Mariakani sandstones, Maji-ya-Chumvi sandstones and shales, igneous intrusives). A variable indicating proximity to year-round surface water bodies (streams, rivers, dams) was determined based on each water point location, taking into account that portions of some surface water bodies in Kwale are saline (Chalala et al., 2017). Those water bodies considered to be year-round and fresh were Mukurumudzi River, Mkanda Dam, tributaries of the Ramisi River (upstream of Mkanda Dam), Lower Koromojo Dam, Umba River and Mwena River.

3.3 Analysis

The analysis consisted of four steps. Initially we evaluated whether the eligibility criteria (i.e. ability to match water points located in 2013 with original installation records) introduced any potential bias. We then conducted three forms of survival analysis, all of which assessed time to failure. First, we ran Kaplan-Meier estimations to compare survival distributions for each of the explanatory variables. Second, we ran Cox proportional hazards

models to identify explanatory variables that exhibited a significant association with the instantaneous risk of failure. Third, we ran accelerated failure time (AFT) models to identify explanatory variables that exhibited a significant association with a water point's operational lifespan. All statistical analysis was performed using statistical software package SPSS (version 24), with the exception of AFT models, which required STATA (version 14.2).

To assess the possible effect of attrition bias, differences in key parameters between matched and unmatched water points were examined by way of independent samples t-tests for continuous variables and Pearson chi-square test for dichotomous variables. Because the precise location of unmatched water points was unknown, the variables assessed were limited to those documented in the installation records and were therefore known for both matched and unmatched water points (i.e. static water level, year of installation, electrical conductivity, yield, user group, season of installation).

Kaplan-Meier estimate analysis was carried out to plot survival distributions for dichotomised explanatory variables, and a series of log rank (Mantel-Cox) tests were run to determine whether any variable led to significantly different survival distributions. Kaplan-Meier estimation is non-parametric, and hence makes no assumption about the shape of the survival function. The cut-off values for static water level, distance to spare parts and electrical conductivity were guided by measures of centrality (mean and median), while the dichotomised threshold for yield was decided with reference to a yield expected to support a handpump supply (0.3 l/s) (MacDonald et al., 2012).

Cox proportional hazards (PH) regression was conducted to ascertain unadjusted and adjusted hazard ratios for the explanatory variables. In the context of this study, hazard ratios represented the change in hazard rate (the instantaneous probability of a water point failing at

any point of time) associated with a unit change in the explanatory variable. The Cox PH model is semi-parametric as no assumptions are made about the shape of the baseline hazard. The effect of an explanatory variable on the hazard rate in a Cox PH model is multiplicative. Nine explanatory variables were included in the analysis: six as categorical variables (locality, user group, surface water, yield, geology, season of installation) and three as continuous variables (spare parts, electrical conductivity, static water level). As a substantial proportion of water points lacked electrical conductivity measurements, two models were run: one with and one without electrical conductivity. A test of non-zero slope based on Schoenfeld residuals was run on both models to ensure compliance with the proportional hazards assumption that underpins Cox models (Grambsch & Therneau, 1994). Accelerated failure time (AFT) models were run (again, one with and one without electrical conductivity), with the choice of survival distribution based on the lowest Akaike information criterion (AIC) value. In contrast to Cox PH models, AFT models are parametric, and thus the shape of the data must be specified. The explanatory variables accelerate or decelerate the subject's time to event. Unadjusted and adjusted time ratios were calculated, representing the proportional difference in survival time associated with a unit change in the explanatory variable. The AFT models included the same explanatory variables as the Cox PH regression.

4. RESULTS

Of the 337 matched water points, 36% were deemed to have failed. The same failure rate was also observed for the 113 KWSP water points that were located in 2013 but could not be matched with a specific entry in the installation records. The failure rate for the other 109 unmatched water points could not be ascertained. Comparisons between matched and unmatched water points revealed significant differences in year of installation, electrical conductivity, and user group (Table 2). On average, matched water points were younger,

produced water with lower electrical conductivity, and were more likely to be located within an institution. There were no significance differences between matched and unmatched water point based on static water level, yield or season of installation. Nor were there any significant associations when including all variables into a multivariable logistic regression model with matched/unmatched status as the outcome variable.

Table 2. Characteristics of water points equipped with Afridev handpumps included in Kwale County Government installation records by whether or not they could be matched with an extant water point in 2013

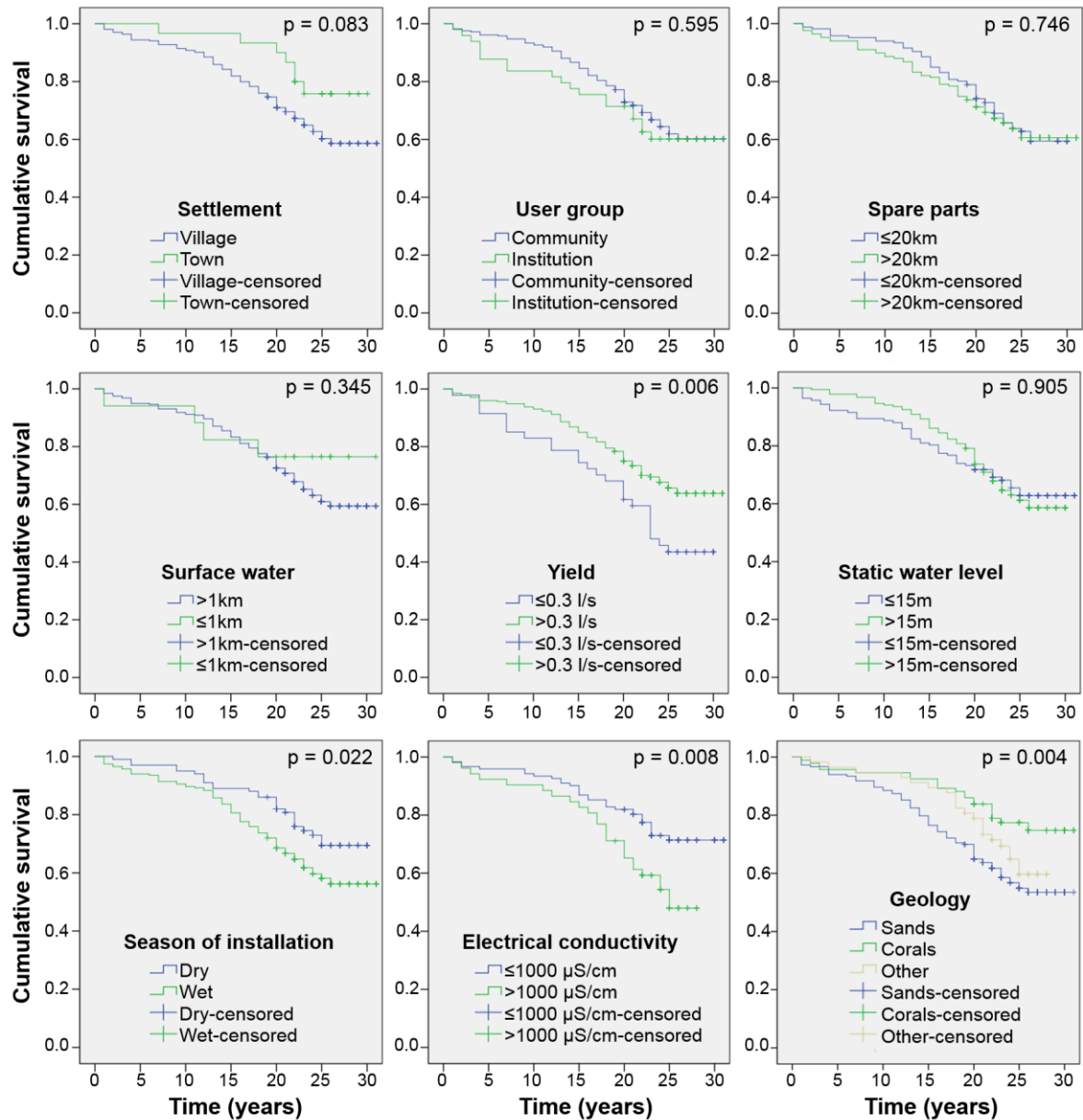
	Matched			Unmatched			p-value ^a
	n	Mean	SD	n	Mean	SD	
Static water level (mbgl)	332	16.6	8.8	211	17.7	10.0	0.194
Installation year	337	1989.3	3.2	222	1988.7	2.9	0.017
Electrical conductivity (µS/cm)	174	987	998	116	1263	1252	0.048
Yield (l/s)	323	0.74	0.51	197	0.71	0.53	0.581
User group (community = 0, institution = 1)	337	0.15	0.36	222	0.08	0.27	0.010
Season (dry = 0, wet = 1)	335	0.70	0.46	218	0.62	0.49	0.068

Note: Bold text indicates statistical significance (p<0.05).

^a Independent samples t-test (with Welch's correction for unequal variances where required) for continuous variables; Pearson Chi-square test for categorical variables.

Kaplan-Meier estimates are presented in Figure 2, as well as Figure S3 and Table S3 in Supplementary Material. Four of the explanatory variables resulted in significantly different survival distributions, namely season of installation, yield, geology, and electrical conductivity.

Figure 2. Kaplan-Meier estimates of the survival functions for Afridev handpumps in Kwale



Note: p-values based on a log rank (Mantel-Cox) test

Results of the Cox proportional hazards models are summarised in Table 4 (see also Figure S4 in Supplementary Material). The multivariable models complied with the proportional hazards assumption both globally and for each individual covariate. In model 1, adjusted

333 hazard rates were significantly higher when electrical conductivity levels were elevated,
334 groundwater was pumped from an unconsolidated sand aquifer, water points were located
335 within an institution, and boreholes were drilled in the wet season. In model 2, adjusted
336 hazard rates were significantly higher when static water level was deeper, groundwater was
337 pumped from an unconsolidated sand aquifer, water points were located further away from
338 spare part suppliers, and boreholes were drilled in the wet season.

339

340

341 Table 4. Results of Cox proportional hazards models for Afridev handpumps in Kwale

Explanatory variable	Dummy coding / Units	Univariable (unadjusted)		Multivariable (adjusted) – Model 1 ^a		Multivariable (adjusted) – Model 2 ^b	
		Hazard Ratio (95% CI)	p-value	Hazard Ratio (95% CI)	p-value	Hazard Ratio (95% CI)	p-value
Settlement	0 = village, 1 = town	0.52 (0.24,1.11)	0.092	0.63 (0.14,2.91)	0.558	0.76 (0.33,1.78)	0.533
User group	0 = community, 1 = institution	1.14 (0.70,1.86)	0.599	2.38 (1.22,4.65)	0.011	1.33 (0.80,2.19)	0.267
Spare parts	kilometres	1.00 (0.99,1.02)	0.440	1.01 (0.98,1.04)	0.437	1.02 (1.00,1.04)	0.040
Surface water	0 = ≤1km, 1 = >1km	1.60 (0.59,4.34)	0.355	0.80 (0.27,2.40)	0.694	1.37 (0.49,3.83)	0.544
Yield	0 = ≤0.3l/s, 1 = >0.3l/s	0.55 (0.35,0.85)	0.007	0.73 (0.39,1.36)	0.322	0.66 (0.42,1.04)	0.075
Static water level	metres below ground level	1.02 (1.00,1.04)	0.109	1.03 (1.00,1.08)	0.079	1.03 (1.00,1.05)	0.028
Season	0 = dry, 1 = wet	1.63 (1.06,2.51)	0.025	2.77 (1.32,5.81)	0.007	2.01 (1.27,3.17)	0.003
Electrical conductivity	100µS/cm	1.02 (1.00,1.04)	0.060	1.03 (1.00,1.05)	0.034		
Geology			0.006		0.033		0.012
Corals vs Sands	0 = sands, 1 = corals	0.46 (0.28,0.74)	0.002	0.46 (0.19,1.09)	0.077	0.59 (0.35,0.99)	0.047
Sands vs Other ^c	0 = other, 1 = sands	1.33 (0.82,2.16)	0.245	2.57 (1.07,6.18)	0.036	2.08 (1.15,3.75)	0.015
Corals vs Other ^c	0 = other, 1 = corals	0.61 (0.33,1.12)	0.111	1.18 (0.37,3.77)	0.782	1.22 (0.58,2.57 4)	0.599

342 Note: Bold text indicates statistical significance (p<0.05). hazard ratio > 1 indicates a higher risk of failure, and a hazard ratio < 1 indicates a lower risk of
343 failure.

344 ^a Model 1 (including EC as explanatory variable) comprises 169 water points

345 ^b Model 2 (excluding EC as explanatory variable) comprises 319 water points

346 ^c ‘Other’ largely comprised of sandstones and shales

347

348
349
350
351
352
353
354
355
356

357

358

Results from the AFT models are presented in Table 5. Weibull models were selected by virtue of their low AIC values, signifying the best fitting model. Variables associated with survival times were largely consistent with findings from the Cox proportional hazards models. In AFT model 3, shorter survival times were significantly associated with higher electrical conductivity, institutional water points, and installation during wet season (all else held constant). In AFT model 4, water points failed earlier when they had deeper static water levels, were underlain by unconsolidated sands, and were installed during wet season (all else held constant).

359 Table 5. Results of accelerated failure time models for Afridev handpumps in Kwale

Explanatory variable	Dummy coding / Units	Univariable (unadjusted)		Multivariable (adjusted) – Model 3 ^a		Multivariable (adjusted) – Model 4 ^b	
		Time Ratio (95% CI)	p-value	Time Ratio (95% CI)	p-value	Time Ratio (95% CI)	p-value
Settlement	0 = village, 1 = town	1.51 (0.91,2.49)	0.108	1.31 (0.53,3.29)	0.558	1.16 (0.69,1.94)	0.580
User group	0 = community, 1 = institution	0.92 (0.67,1.27)	0.624	0.61 (0.40,0.92)	0.017	0.85 (0.62,1.16)	0.297
Spare parts	kilometres	1.00 (0.99,1.01)	0.455	1.00 (0.98,1.01)	0.615	0.99 (0.98,1.00)	0.051
Surface water	0 = ≤1km, 1 = >1km	0.73 (0.38,1.41)	0.355	1.09 (0.56,2.11)	0.798	0.82 (0.44,1.54)	0.544
Yield	0 = ≤0.3l/s, 1 = >0.3l/s	1.45 (1.09,1.93)	0.010	1.20 (0.82,1.75)	0.359	1.27 (0.96,1.69)	0.091
Static water level	metres below ground level	0.99 (0.97,1.00)	0.088	0.98 (0.96,1.00)	0.104	0.98 (0.97,1.00)	0.028
Season	0 = dry, 1 = wet	0.73 (0.55,0.97)	0.028	0.56 (0.36,0.89)	0.013	0.66 (0.49,0.88)	0.004
Electrical conductivity	100µS/cm	0.99 (0.97,1.00)	0.058	1.02 (1.00,1.03)	0.035		
Geology			0.003		0.060		0.018
Sands vs Corals	0 = corals, 1 = sands	0.60 (0.44, 0.83)	0.002	0.62 (0.37,1.05)	0.074	0.71 (0.52,0.99)	0.042
Sands vs Other ^c	0 = other, 1 = sands	0.85 (0.62,1.16)	0.306	0.62 (0.36,1.04)	0.072	0.66 (0.46,0.95)	0.026
Corals vs Other ^c	0 = other, 1 = corals	1.41 (0.95,2.09)	0.089	1.00 (0.50,1.99)	0.989	0.93 (0.59, 1.46)	0.743

360 Note: Bold text indicates statistical significance (p<0.05). A time ratio > 1 indicates the time to failure is prolonged, and a time ratio < 1 indicates the time to
361 failure is shortened.

362 ^a Model 3 (including EC as explanatory variable) comprises 169 water points. Weibull shape parameter value of $p = 1.67$ (95% CI, 1.30-2.13)

363 ^b Model 4 (excluding EC as explanatory variable) comprises 319 water points. Weibull shape parameter value of $p = 1.63$ (95% CI, 1.38-1.93)

364 ^c ‘Other’ largely comprised of sandstones and shales

365

366

5. DISCUSSION

The results suggest hydrogeological and geographical factors have impinged upon the sustainability of rural water supplies in Kwale over the course of three decades. Water points appear less likely to survive when they supply water with higher electrical conductivity, draw on deeper groundwater, are underlain by unconsolidated sands, and are installed in the wet season. Water points located within an institution and situated further away from spare parts supplies also have higher failure risks, though these associations exhibit less consistency across the models.

The relationship between electrical conductivity and water point failure probably results from user rejection or abandonment of the handpump because of the unsatisfactory taste of the water supplied. For every increase of 100 $\mu\text{S}/\text{cm}$ at the time of installation, the risk of failure rises by 3%, and a water point's operational lifespan is reduced by 2%. A significant relationship between electrical conductivity and palatability has been previously demonstrated in Kwale (Foster and Hope, 2016; see also Fig. S5 in Supplementary data), and has been observed elsewhere in sub-Saharan Africa (Langenegger, 1989). This result points to the importance of monitoring groundwater quality in Kwale, particularly in light of reports that groundwater salinity levels have risen over previous decades, and the ongoing threat of seawater intrusion (Tole, 1997).

Water points with deeper static water levels appeared more prone to failure. With each metre of depth, the models suggested an increase in the hazard rate of 3% (Coxmodel 2) and the end of a water point's operational life is reached 2% earlier (AFT model 4). The relationship between static water level and water point failure may arise due to a higher frequency of breakdowns that would be expected to accompany greater pumping lifts by virtue of increased stresses, greater wear and tear per unit of water produced as a result of lower pump

efficiency, or simply a consequence of a greater number of components (e.g. pipes and rods) increasing the probability of a ‘weak link’. These factors could in turn accelerate the occurrence of major mechanical failure or frequency of routine failures, thereby amplifying the recurrent costs that need to be borne by water users. Early operational data describing the lifespans of Afridev bearings in Kwale provides tentative support for the proposition that mean time between breakdowns is influenced by groundwater depth (Reynolds, 1992). An alternative explanation is that deeper boreholes are more complex to drill and construction quality risks are greater. It should be noted that the association observed here is specific to boreholes, and may not apply to hand-dug wells. Despite their shallower water level, there is evidence that hand-dug wells have a higher failure rate than boreholes (Foster, 2013).³

The higher failure risk exhibited by water points underlain by Pliocene and Pleistocene sand formations concurs with anecdotal evidence that these aquifers have caused boreholes to backfill with fine sands. Logically, fine sands may also accelerate wear and tear of the handpump by impairing pump efficiency and abrading rubber components. Further investigation is needed to assess whether other hydrochemical or hydraulic mechanisms underlie the link between geology and water point lifespans, noting that differences in yield, static water level and electrical conductivity were controlled for in the Cox PH and AFT models. For example, in some areas of Kwale low pH has been known to cause rapid corrosion of mild steel pump rods (Reynolds, 1992), while hardness also poses problems for some boreholes (Thambu, 1989).

The link between the risk of failure and season at time of drilling completion is consistent the findings of Harvey (2001), who assessed predictors of rapid-onset borehole failure in Ghana.

³ This is perhaps due to a greater propensity to dry up, a large diameter that can allow communities to continue fetching water with a rope and bucket in the event of a handpump breakdown, and proximate alternative groundwater sources that can be developed at low cost

The author in that study concluded that the relationship between borehole failure and rainfall was due to drillers failing to adequately take into account seasonal water level fluctuations, as evidenced by the constancy of cylinder depth relative to dynamic water level across the seasons. Unfortunately cylinder depths were not systematically documented in the Kwale installation records, hence it is not possible to confirm or rule out whether similar issues underlie the emergent seasonal pattern. However, we believe this is unlikely to be the primary explanation for the association in the Kwale context, as seasonal groundwater fluctuations tend to be relatively modest. Though there has been no long-term monitoring of groundwater levels, observations from the time of early installations in 1983 suggest a 1.0 to 2.5 metre difference in static water level between wet and dry season in the Pleistocene corals (Thambu, 1987). In comparison, data presented by Reynolds (1992) suggest 19 handpumps installed under the KWSP had cylinder settings that were on average nine metres deeper than the static water level.

A more plausible explanation for the multivariable association in Kwale is that season of installation acts as a control variable for hydrogeological covariates. This proposition is supported by previous investigations in Kwale that have shown seasonal variations in both static water level and electrical conductivity (Thambu 1989, Mzuga et al., 1998). In other words, the results point to the importance of analysis adjusting for the season in which hydrogeological variables are measured.

The association between water point survival and distance to spare parts conforms to commonly held views about supply chain viability (Harvey & Reed, 2006), and agrees with previous findings about the relationship between functionality and proximity to spare parts in Sierra Leone (Foster, 2013). The Cox model 2 results indicate a 2% increase in the hazard rate for every additional kilometre between the water point and spare parts retailer. In Kwale, the location of spare part suppliers is widely known and we expect the relationship observed

is not due to a lack of awareness per se, but rather that those communities farthest away incur the highest transaction costs and greatest logistical challenges to procure spare parts (noting that some communities may procure parts themselves, while others rely on a pump mechanic to source the parts). The relationship may also be compounded or confounded by the possibility that the more remote communities have had higher levels of poverty and a lower capacity to pay for ongoing maintenance costs.

There is little contextual information to pinpoint why institutional water points exhibited higher hazard rates relative to community water points. There may be inherent differences related to usage levels, management and financing that could all play a role. The situation is somewhat complicated by the fact that delineation between community and institutional roles and responsibilities is not always well defined. The association must also be interpreted with caution due to the possible effect of attrition bias. Institutional water points were easier to match than community water points because a name referring to a school, police station or health centre tends to be unique. As a result, a disproportionate number of failed institutional water points appear to have been included in the matched sample (as evidenced in Table 2).

Overall, the shape of the survival functions (see Figures S3) suggests a relatively high proportion of water points survive the first ten years (~90%), followed by a downward steeping of the curve such that the survival rate drops to around 60% by year 25. This is further illustrated by the AFT models, which have a Weibull shape parameter value of $p = 1.7$, indicating a monotonically increasing hazard rate over time. At first glance, this seems to contradict national water point datasets from elsewhere which suggest a high failure rate in the early years, followed by a flat-lining beyond the tenth year (Carter & Ross, 2016). This apparent incongruence may arise for several reasons. The monotonically increasing hazard rate observed in the Kwale data may reflect a plausible pattern whereby failure risks grow as waterpoints age. Over time, boreholes and handpumps may be more likely to experience

failure modes that are technically and financially harder to repair, or alternatively the breakdowns become more frequent and give rise to higher (and at times prohibitive) running costs. Managerial structures may also weaken over time: committees may become inactive or dissolve entirely or changes in committee membership may diminish management capabilities as new members lack the training and know-how imparted at the time of implementation.

The shape of the survival functions seemingly contradict national water point datasets from other sub-Saharan African countries, which instead tend to show a relatively high year-on-year failure rate in the early years after installation, followed by a flat-lining beyond the tenth year (Carter and Ross, 2016). This apparent incongruity may be linked to a bias that Carter and Ross (2016) dub the “denominator problem”. In the case of cross-sectional water point datasets, failure rates may appear to plateau after a certain age because some older abandoned water points have disappeared and are excluded from the analytical frame, leading to an overestimation of the functionality rate for this older cohort (Carter and Ross, 2016; Jiménez and Pérez-Foguet, 2011). The water point survival rate in the early years of this retrospective assessment in Kwale could be overestimated for the same reason, potentially causing the survival functions to be upwardly biased. The observation that unmatched water points were slightly older than matched water points (and thus more likely to have failed) provides tentative support for this hypothesis.

Putting in place operation and maintenance arrangements that counter the risk factors observed in this study will be vital to achieving universal access to safe drinking water in Kwale, and in rural sub-Saharan Africa more generally. With a majority of matched water points functioning 25 years after installation, the results demonstrate that with adequate maintenance handpumps can achieve impressive longevity. However, this still falls short of what will be needed to achieve a safe water Sustainable Development Goal (SDG) that is

premiered upon universality. Provision of external support to communities is now widely advocated as a way to improve the sustainability of rural water services in sub-Saharan Africa. The results from this investigation suggest these external support mechanisms will need to mitigate the heterogeneous environmental and geographical obstacles that water users encounter. This may require providing extra or tailored financial and technical support for those communities that face conditions less conducive to sustainable operation and maintenance, or perhaps consideration of alternative techno-institutional approaches to rural water service delivery. Heuristics based upon the predictors observed in this and other similar investigations may provide a guiding framework as to where and how support could be differentiated. More simply, a centralised maintenance service could facilitate cross subsidies between water points, eliminating inherent financial and technical disadvantages experienced by some communities (Hope et al., 2012). Examples of this approach have been documented elsewhere in Kenya (Thomson & Koehler, 2016; Goodall & Katilu, 2016; Foster & McSorley, 2016).

In addition to usual caveats regarding omitted variable bias, this study has several limitations. First, it is possible that some water points may have previously been non-functional for more than a year but subsequently rehabilitated (in which case, the time to event is biased upwards). Second, the duration of downtime for non-functional handpumps was reliant on self-reported estimates by water users, and therefore susceptible to recall bias. Third, limited information was available on the processes and quality control methods undertaken by borehole drillers and program implementers when measuring and documenting water point characteristics. In recognition of these weaknesses, we would stress the continued importance of high quality and rigorous research designs that are capable of unravelling causal pathways in a more nuanced way.

Notwithstanding these limitations, the study demonstrates the amenability and advantages of applying survival analysis to water point sustainability evaluations, particularly in tandem with installation data. First, the explanatory variables utilised are exogenous in nature and pertained to the time of installation. In addition to avoiding concerns relating to reverse causation, this means insights are highly relevant to practitioners involved in programme implementation. Second, the outcome variable provides information about the longer-term survival time of the handpump rather than a simple binary functionality status indicator. This is arguably more useful to policymakers and practitioners as it better reflects the return on the initial infrastructure investment. Third, as groundwater characteristics used in the study were captured at the time boreholes were drilled, associations with water quality parameters were able to be assessed specifically for each water point in a low-cost fashion, irrespective of subsequent functionality status. As such, the approach is well suited to assessing water point failure risks in resource-constrained contexts.

6. CONCLUSION

Survival analysis provides an instructive approach to estimating failure risks for rural water points, and its application to handpump-equipped boreholes in Kwale suggests that groundwater depth, water quality, geology and spare part supply chains are important determinants of sustainability. The season in which groundwater characteristics are measured is also an important factor that should be controlled for in multivariable analysis of water point failure. The findings are of broad relevance given the preponderance of community-managed handpumps across rural sub-Saharan Africa. The importance of insights is strengthened by the application of a novel analytical technique that avoids some of the shortcomings of logistic regression modelling, as well as the multi-decadal nature of the data underpinning the study. Future investigations should look to incorporate a broader range of hydrogeological and socio-economic variables where possible, including time-series data for

time varying characteristics, and triangulate results with a more detailed understanding of failure modes and social processes to explain causal routes. This information will be critical for policymakers and practitioners to facilitate more sustainable water services and achieve a greater return on infrastructure investments.

ACKNOWLEDGEMENTS

This paper is an output from the “New Mobile Citizens and Water point Sustainability in Rural Africa” (ES/J018120/1) and “Groundwater Risk Management for Growth and Development” (NE/M008894/1) research projects supported by the UK Department for International Development (DFID), the UK Economic and Social Research Council and the UK Natural Environment Research Council. The views expressed in this paper are those of the authors and do not necessarily represent those of DFID. The authors would like to acknowledge Rural Focus Ltd and the Kwale County Government for their assistance with data collection.

REFERENCES

- Baumann, E.; Furey, S. *How Three Handpumps Revolutionised Rural Water Supplies: A Brief History of the India Mark II/III, Afridev and the Zimbabwe Bush Pump*; Rural Water Supply Network: St. Gallen, 2013.
- Bonsor, H.; Oates, N.; Chilton, P.; Carter, R.; Casey, V.; MacDonald, A.; Calow, R.; Wilson, P.; Tumutungire, M.; Bennie, M. *A hidden crisis: strengthening the evidence base on the*

560 *sustainability of rural groundwater supplies: results from a pilot study in Uganda*; British
561 Geological Survey: Keyworth, 2015.

562 Bradburn, M.; Clark, T.; Love, S.; Altman, D. Survival analysis part II: multivariate data
563 analysis – an introduction to concepts and methods. *Brit. J. Cancer* **2003**, 89(3), 431-436.

564 Carter, R; Ross, I. Beyond ‘functionality’ of handpump-supplied rural water services in
565 developing countries. *Waterlines* **2016**, 35 (1), 94-110.

566 Caswell, P. *Geology of the Mombasa-Kwale Area*. Report No. 24. Geological Survey Kenya:
567 Nairobi, 1953.

568 Chalala, A.; Chimbevo, L.; Kahindo, J.; Awadh, M.; Malala, J. Sea water intrusion and
569 surface water salinity and its influence on irrigation water quality in Ramisi Area, Kenya
570 **2017**, 12(1), 1-13.

571 Clark, T.; Bradburn, M.; Love, S.; Altman, D. Survival analysis part I: basic concepts and
572 first analyses. *Brit. J. Cancer* **2003**, 89 (2), 232-238.

573 Cox, D. Regression Models and Life-Tables. *J. Roy. Stat. Soc. B. Met.* **1972**, 34 (2), 187–220.

574 Cronk R.; Bartram, J. Factors influencing water system functionality in Nigeria and
575 Tanzania: a regression and Bayesian network analysis. *Env. Sci. Tech.* **201**, **51** (19), 11336–
576 **11345**.

577 Fisher, M.; Shields, K.; Chan, T.; Christenson, E.; Cronk, R.; Leker, H.; Samani, D.; Apoya,
578 P.; Lutz, A.; Bartram, J. Understanding handpump sustainability: Determinants of rural water
579 source functionality in the greater Afram plains region of Ghana. *Water Resour. Res.* **2015**,
580 51(10), 8431-8449.

581 Foster, T. Predictors of sustainability for community-managed handpumps in Sub-Saharan
582 Africa: Evidence from Liberia, Sierra Leone, and Uganda. *Env. Sci. Tech.* **2013**, 47 (21),
583 12037-12046.

584 Foster, T.; Hope, R. A multi-decadal and social-ecological systems analysis of community
585 water point payment behaviours in rural Kenya. *J. Rur. Stud.* **2016**, 47, 85-96.

586 Foster, T.; Hope, R. Evaluating waterpoint sustainability and access implications of revenue
587 collection approaches in rural Kenya. *Water Resour. Res.* **2017**, 53 (2), 1473-1490.

588 Foster, T.; McSorley, B. *An Evaluation of the BluePump in Kenya and The Gambia*; ISF-
589 UTS & Oxfam, **2016**.

590 Goodall, S.; Katilu, A. FundiFix: exploring a new model for maintenance of rural water
591 supplies. *39th WEDC International Conference: Ensuring availability and sustainable*
592 *management of water and sanitation for all*, Kumasi, Ghana, **2016**.

593 Grambsch, P.; Therneau, T. Proportional hazards tests and diagnostics based on weighted
594 residuals. *Biometrika* **1994**, 81 (3), 515-526.

595 Harvey, P. Borehole Sustainability in Rural Africa: An analysis of routine field data. *30th*
596 *WEDC International Conference: People-centred Approaches to Water and Environmental*
597 *Sanitation*, Vientiane, Laos PDR, **2001**.

598 Harvey, P.; Reed, R. Sustainable supply chains for rural water supplies in Africa. *P. I. Civil*
599 *Eng.-Eng. Su.* **2006**, 159 (1), 31-39.

600 Hope, R.; Foster, T.; Thomson, P. Reducing risks to rural water security in Africa. *Ambio*
601 **2012**, 41(7), 773-776.

602 ILRI, *The geology of Kenya*, **2007**. Available at
603 <http://192.156.137.110/gis/search.asp?id=478>. (accessed July 10, 2017).

604 Jiménez, A.; Pérez-Foguet, A. Water point mapping for the analysis of rural water supply
605 plans: case study from Tanzania. *J. Water Resour. Plann. Manage.*, **2011**, 137 (5), 439-447.

606 Kaplan, E.; Meier, P. Nonparametric estimation from incomplete observations. *J. Am. Stat.*
607 *Assoc.* **1958**, 53 (282), 457-481.

608 Langenegger, O. Groundwater Quality-An Important Factor for Selecting Handpumps. *Dev.*
609 *Water Sci.* **1989**, 39, 531-541.

610 MacArthur, J. *Handpump Standardisation in Sub-Saharan Africa: Seeking a champion*; Rural
611 Water Supply Network: St Gallen, **2015**.

612 MacDonald, A.; Bonsor, H.; Dochartaigh, B.; Taylor, R. Quantitative maps of groundwater
613 resources in Africa. *Environmental Research Letters* **2012**, 7 (2), 024009.

614 McCommon, C.; Warner, D.; Yohalem, D. *Community Management of Rural Water Supply*
615 *and Sanitation Services*; UNDP-World Bank Water and Sanitation Program, Washington,
616 DC, **1990**.

617 Mzuga, J. M., Tole, M. P., & Ucauwun, E. K. The impact of geology and pit latrines on
618 groundwater quality in Kwale District. *Coast. Ecol. Ser.* **1998**, 4, 85-96.

619 Narayan-Parker, D. *People, Pumps and Agencies: The South Coast Handpump Project*;
620 PROWESS/UNDP: Nairobi, **1988**.

621 Reynolds, J. *Handpumps: Toward a Sustainable Technology*; UNDP-World Bank Water and
622 Sanitation Program, Washington DC, **1992**.

623 RWSN, *Handpump Data, Selected Countries in Sub-Saharan Africa*, 2009. <http://www.rural->
624 [water-supply.net/_ressources/documents/default/203.pdf](http://www.rural-water-supply.net/_ressources/documents/default/203.pdf). (accessed September 19, 2016).

625 Thambu, G. *Variations in quantity and quality of groundwater in the South Coast Region of*
626 *Kenya*. Masters Dissertation, University of Manitoba, **1987**.

627 Thambu, G. Conditions in shallow wells in the South Coast Region of Kenya. *Water Int.*
628 **1989**, 14 (3), 112-121.

629 Thomson, P.; Hope, R.; Foster, T. Is silence golden? Of mobiles, monitoring, and rural water
630 supplies. *Waterlines* **2012**, 31 (4), 280-292.

631 Thomson, P.; Koehler, J. Performance-oriented monitoring for the water SDG—challenges,
632 tensions and opportunities. *Aquatic Procedia* **2016**, 6, 87-95.

633 Tole, M. Pollution of Groundwater in the Coastal Kwale District, Kenya. *Sustainability of*
634 *Water Resources under Increasing Uncertainty: Proceedings of the Rabat Symposium*. Rabat,
635 **1997**.

636 Walters, J.; Javernick-Will, A. Long-term functionality of rural water services in developing
637 countries: a system dynamics approach to understanding the dynamic interaction of factors.
638 *Env. Sci. Tech.* **2015**, 49 (8), 5035-5043.

639 Whittington, D.; Davis, J.; Prokopy, L.; Komives, K.; Thorsten, R.; Lukacs, H.; Bakalian, A.;
640 Wakeman, W. How well is the demand-driven, community management model for rural
641 water supply systems doing? Evidence from Bolivia, Peru and Ghana. *Water Policy* **2009**, 11
642 (6), 696–718.

643

Supplementary Material

Risk factors associated with rural water supply failure: a 30-year retrospective study of handpumps on the south coast of Kenya

Tim Foster, Juliet Willetts, Jacob Katuva, Patrick Thomson, Rob Hope

Figure S1. Afridev handpump installations: 1983-1995 (Kwale County Government records)

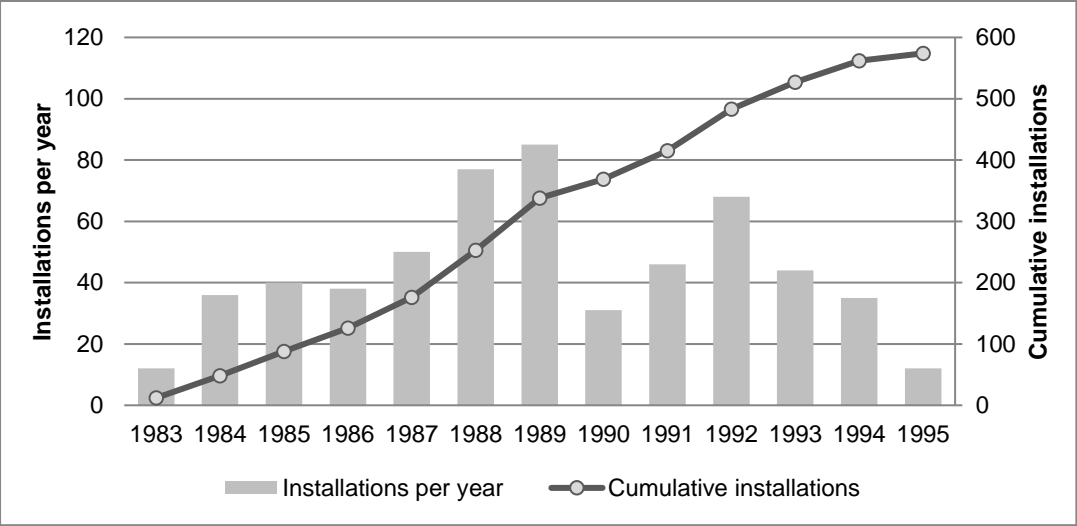


Figure S2. Geological map of (A) Kenya and (B) study site in Kwale

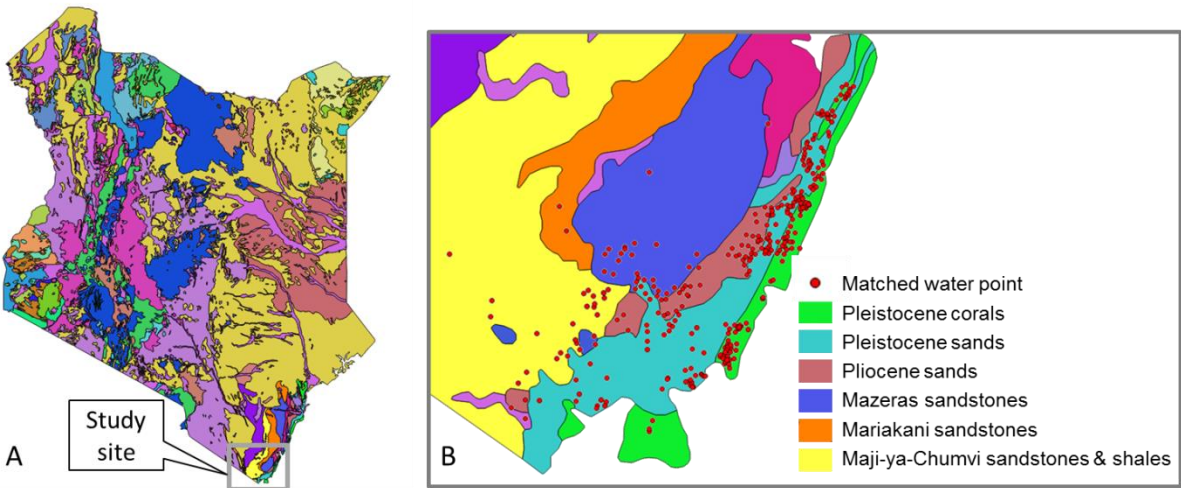


Figure S3. Kaplan-Meier estimate of the survival function for Afridev handpumps in Kwale

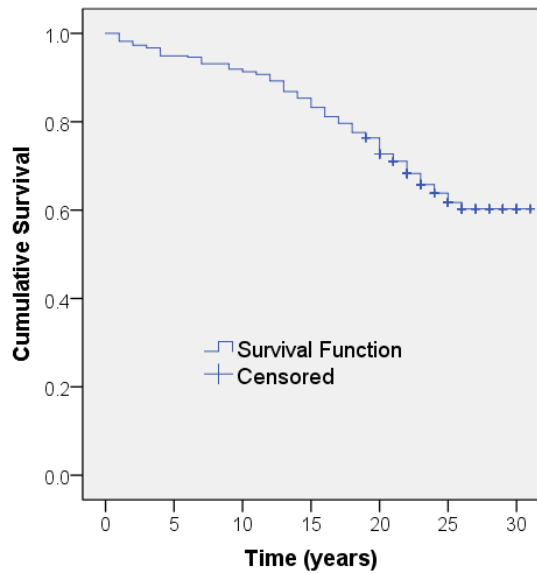


Figure S4. Cox proportional hazards survival functions at mean of covariates for Afridev handpumps in Kwale

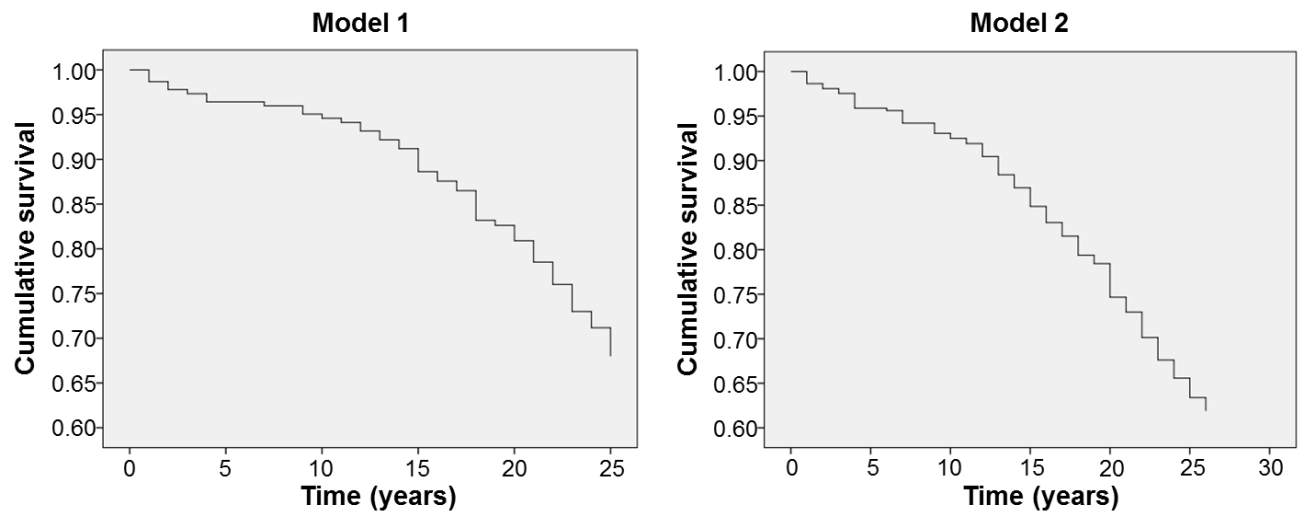


Figure S5. Proportion of respondents Kwale describing taste of water as ‘good’ in both dry and wet season (n=2,142).

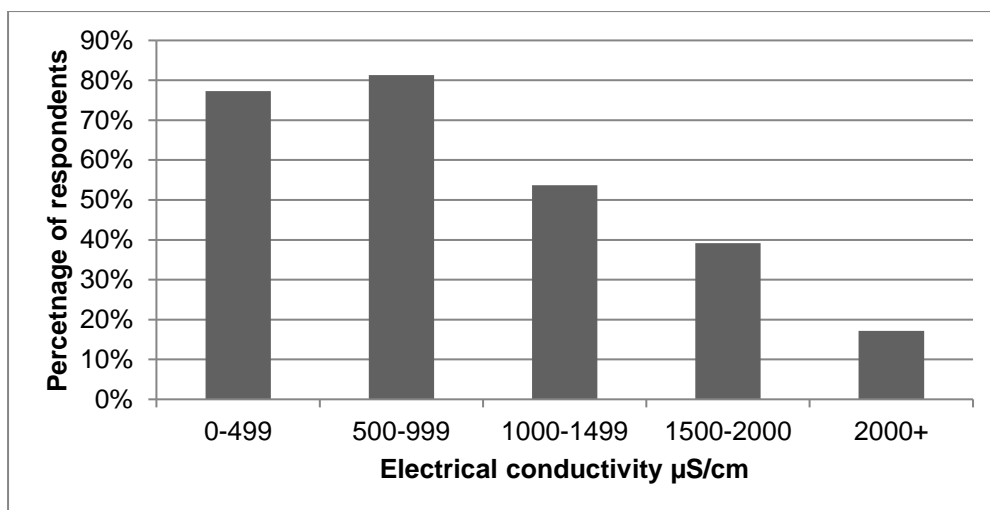


Table S1. Location of water points equipped with Afridev handpumps included in Kwale County Government installation records, including the number of water points in the installation records that could be matched with extant water points in 2013

Area	No. water points	No. matched water points	No. unmatched water points	% matched
Diani	93	55	38	59.1
Kikoneni	70	35	35	50.0
Kinondo	75	57	18	76.0
Lukore	15	10	5	66.7
Lunga Lunga	16	8	8	50.0
Majimboni	4	2	2	50.0
Mkongani	5	4	1	80.0
Msambweni	155	93	62	60.0
Mwaluphamba	3	1	2	33.3
Mwereni	6	3	3	50.0
N'gombeni	14	11	3	78.6
Pongwe Kidumu	27	16	11	59.3
Tiwi	44	24	20	54.5
Tasimba	3	1	2	33.3
Waa	29	17	12	58.6
Total	559	337	222	60.3

669 Table S2. Explanatory variables included in the study, including data sources and justifications

Explanatory variable	Categories	Data source	Hypothesis
Hydrogeological			
Static water level (mbgl)		Government installation records	Greater depth increases risk of failure (greater stresses, accelerated wear and tear)
Yield (l/s)		Government installation records	Lower yield increases risk of failure (dissatisfaction with discharge, greater drawdown increases wear and tear)
Geology	1: Sands 2: Corals 3: Other	Classifications based on Caswell (1953), using digital map available at ILRI (2007)	Unconsolidated sands increase risk of failure (backfilling borehole, abrading pump components)
Season of installation	0: Dry season (Jan-Apr) 1: Wet season (May-Dec)	Government installation records	Rainfall season affects hydrogeological parameters (e.g. static water level for a given water point likely to be deeper in dry season than wet season)
Electrical conductivity (µS/cm)		Government installation records	Higher EC increases risk of failure (users reject saline water)
Geographical			
Surface water	0: fresh and year-round surface water within 1km 1: no fresh and year-round surface water within 1km	Calculated based on water point location relative to year-round fresh surface water bodies	Closer year-round fresh surface water increases risk of failure because availability of alternative water sources reduces incentives to repair a waterpoint
Spare parts (km)		Calculated based on water point location relative to Ukunda town	More distant spare part suppliers increase risk of failure (due to higher transaction costs of buying parts)
Locality	0: village 1: town	Designation of 'Town' based on official definition of Kenya National Bureau of Statistics. Town boundaries determined according those presented in http://majidata.go.ke	Densely populated communities might increase failure risk (due to alternative water sources, such as piped schemes and vendors) or reduce failure risk (because of higher socio-economic status of users and greater access to external support)
User group	0: community 1: institution	Government installation records	Institutional water points could either increase failure risk (if maintenance funds are more limited and/or usage is heavier) or reduce failure risk (if maintenance funds readily available and/or usage is lighter)

670

671

672 Table S3. Kaplan-Meier estimates for Afridev handpumps in Kwale

	Observations	Events ^a (%)	Censored ^b (%)	Survival time (years)	Log Rank (Mantel-Cox)	
				Mean (95% CI)	χ^2	p-value
All	334	120 (36.9)	214 (64.1)	25.0 (24.1-26.0)		
Settlement						
Town	30	7 (23.3)	23 (76.7)	27.3 (25.4-29.3)	3.00	0.083
Village	304	113 (37.2)	191 (62.8)	24.7 (23.7-25.7)		
User group						
Community	285	101 (35.4)	184 (64.6)	25.2 (24.3-26.2)	0.28	0.595
Institution	49	19 (38.8)	30 (61.2)	23.6 (20.7-26.5)		
Spare parts (km)						
≤20	166	59 (35.5)	107 (64.5)	24.8 (23.6-26.0)	0.11	0.746
>20	167	61 (36.5)	106 (63.5)	24.6 (23.2-26.0)		
Surface water (fresh, year-round)						
≤ 1km	17	4 (23.5)	13 (76.5)	24.9 (24.0-25.9)	0.89	0.345
> 1km	317	116 (36.6)	201 (63.4)	26.2 (21.8-30.5)		
Yield (l/s)						
≤0.3	47	26 (55.3)	21 (44.7)	21.7 (19.1-24.3)	7.67	0.006
>0.3	273	89 (32.6)	184 (67.4)	25.6 (24.7-26.6)		
Static water level (mbgl)						
≤15	142	48 (33.8)	94 (66.2)	24.6 (23.0-26.2)	0.01	0.905
>15	187	70 (37.4)	117 (62.6)	24.9 (23.9-25.9)		
Season of installation						
Dry	100	27 (27.0)	73 (73.0)	26.1 (24.7-27.5)	5.26	0.022
Wet	232	93 (40.1)	139 (59.9)	24.2 (23.0-25.4)		
Geology						
Sands	184	78 (42.4)	106 (57.6)	23.6 (22.3-25.0)	10.9	0.004
Corals	93	21 (22.6)	72 (77.4)	27.3 (25.8-28.8)		
Other ^c	57	21 (36.8)	36 (63.2)	24.0 (22.4-25.7)		
Electrical conductivity (μS/cm)						
≤1000	122	32 (26.2)	90 (73.8)	26.6 (25.2-28.0)	7.06	0.008
>1000	52	25 (48.1)	27 (51.9)	22.0 (19.9-24.1)		

673 Note: Bold text indicates statistically significant association (p < 0.05).

674 ^a ‘Events’ defined as those water points that had failed (non-functional for more than one year). ^b ‘Censored’675 defined as those water points that had not yet failed. ^c ‘Other’ geology comprises sandstones and shales

676