Migration and Mortality: Identifying Selection and Exposure Effects in Local Populations using Longitudinal Data

Authors: *Philippe Bocquier[1,2], Carren Ginsburg[2], Donatien Béguy[3], Sulaimon Afolabi[2], Orvalho Augusto[4], Karim Derra[5], Frank Odhiambo[6,12], Mark Otiende[7], Abdramane Soura[8], Pascal Zabré[9], Michael J. White[10, 2], Mark A. Collinson[2, 11,12]*

Author Affiliations:

[1] Centre de Recherches en Démographie, Université Catholique de Louvain, Louvain-la-Neuve, Belgium

[2] Medical Research Council/Wits Rural Public Health and Health Transitions Research Unit

(Agincourt), School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

[3] African Population and Health Research Centre, Nairobi, Kenya

[4] Manhiça HDSS, Manhiça, Mozambique

[5] Clinical Research Unit of Nanoro (CRUN), Nanoro, Burkina Faso

[6] KEMRI & CDC - Centre for Global Health Research, Kisumu, Kenya

[7] Kenya Medical Research Institute, Wellcome Trust Research Programme, Kilifi, Kenya

[8] Institut Supérieur des Sciences de la Population (ISSP), Université de Ouagadougou, Burkina Faso

[9] Centre de Recherche en Santé de Nouna, Nouna, Burkina Faso

[10] Population Studies and Training Centre, Brown University, USA

[11] Umeå Centre for Global Health Research, Umeå University, Sweden

[12] INDEPTH Network, Accra, Ghana

Keywords: Internal Migration, sub-Saharan Africa, Mortality, Health and Demographic Surveillance System

Abstract:

Migration has been hypothesised to be selective on health but this healthy migrant hypothesis has generally been tested at destinations, and for only one type of flow, from deprived to better-off areas. The circulatory nature of migration is rarely accounted for. This study examines the relationship between different types of internal migration and mortality in Health and Demographic Surveillance System (HDSS) populations in West, East, and Southern Africa, and asks how the processes of selection and exposure explain the migration-mortality relationship experienced in these contexts. The paper uses longitudinal data representing approximately 900 000 people living in nine sub-Saharan African HDSS sites of the INDEPTH Network. Event History Analysis techniques are employed to examine the relationship between all-cause mortality and migration status, over periods ranging from 3 to 14 years. The study confirms the importance of migration in explaining variation in mortality, and the diversity of the migration-mortality relationship over a range of rural and urban local areas in the three African regions. The results confirm that the pattern of migration-mortality relationship is mainly generated by the combination of two processes: selection and exposure. However, the "healthy migrant" and the "unhealthy return migrant" hypotheses are not universal and sometimes contradicted. Consequences for public health policy are drawn.

1. Introduction

Migration, understood as a change in usual residence, is a much more common event than other demographic phenomena such as death or birth, especially with secular fertility decline. In the mid-2000s, the median Aggregate Crude Migration Intensity (ACMI), a measure of all permanent changes of address within a country, was 7.5% for 1-year period in a range of 45 countries around the world (Bell et al. 2015), and this is not even accounting for international migrations. To give a sense of scale, the world crude birth rate was only 2% a year and the world crude death rate less than 1% (United Nations Population Division 2014). Across the 45 selected countries for which data were available, only four had an 1-year ACMI smaller than 2% (Bell et al. 2015).

Migration is not only a major demographic event, but it also has the potential to influence the other demographic dimensions. In this article, we investigate the relationship between migration and health, using mortality as major indicator of health. This relationship is important because health status may both impede and stimulate migration, while migration, often motivated by economic benefits, can result in negative health outcomes, possibly leading to death (Gerritsen et al. 2013b). Although studies abound on mortality and to a lesser extent on migration, their relationship has been far less investigated. In this paper, we revisit the theory about this relationship, accounting for different types of migration flow and the level of health risk in origin and destination areas. We design a method to identify a set of hypotheses attached to the migration-mortality relationship, and apply this method to interpreting data collected in local populations in West, East, and Southern Africa that present different patterns of mortality and migration. By identifying the most likely explanation for the divergent patterns seen in these different settings, we aim to better distinguish the categories of migrants for health interventions.

2. Literature review

Although the migration-mortality is not foreign to the broader issue of epidemiological transition, this section will not attempt to review the role of migration in this transition, but will rather build on previous reviews on the subject (Collinson et al. 2014). In the context of low- and middle-income countries (LMICs), the health transition has been at the same time spectacular in its speed and more heterogeneous than in higher-income countries (HICs) (Salomon and Murray 2002). The transition has led to a general decline in mortality but also, particularly in African countries, to a double burden of disease characterised by the emergence of non-communicable diseases (NCDs) and life-style diseases associated with urbanisation (Ezzati et al. 2005), which coexist with persistent, new and revitalised diseases, such as malaria, HIV/AIDS and TB (Boutayeb 2006). This double burden runs counter to mortality decline. Whereas urbanisation has generally contributed positively to health in the past, there are concerns that under conditions of slow economic development and weak infrastructure management it could actually drive an increase in cardio-vascular disease (CVD) (Yusuf et al. 2001) as well as in respiratory and diarrheal diseases linked to bad environmental conditions (Harpham 2009).

Migration plays an important role in sustaining livelihoods in LMIC countries. As people migrate, remittances and information circulate and help to maintain links between sending and receiving communities (White and Lindstrom 2006). However migration may improve well-being and at the same time expose migrants to health risks. With respect to mortality, it is not clear whether the net effect of migration is positive or negative, and in which circumstances. Considering the sheer volume of migration and its high sensitivity to livelihood conditions, it is necessary to carefully examine the hypotheses relating to migration and health (for references and glossary of terms used in migrant and health analysis, see Urquia and Gagnon 2011).

The first and most-utilised hypothesis concerning migration and health is the "healthy (im)migrant" hypothesis. This hypothesis proposes that migrants are selected in their place of

origin amongst the more healthy since they must prepare to adapt to their new place of residence. Positive selection on health would then operate through migration. In ordinary migration, because of a high selection effect at origin, the health of migrants can actually be better than that of non-migrants at destination the location. This *selection effect* could explain the epidemiological paradox that even if originating from places with high health risks, migrants may have better health than the non-migrants in destination areas living in superior (health) conditions (Urquia and Gagnon 2011).

This would apply provided that the migration conditions are not too stringent, as may be the case for refugees and internally displaced people. The possible effect of the migration conditions is called the *disruption effect* and is usually attributed to the conditions around the time of migration (just before and just after, the specific time span to be defined by the migration itself). This effect has been particularly studied for reproductive health (Choi 2014; Goldstein 1973; Hervitz 1985; Kulu and Steele 2013).

After some time at the destination, the migrants' health may gradually converge to that of non-migrants. This *adaptation effect* (also named assimilation effect) is observed over time, i.e. with duration of residence, and can only present if there is a difference between the health of non-migrants and the health of migrants just after their arrival. It is often presented as a loss of (negative or positive) selection effect over time (Urquia and Gagnon 2011).

Lastly, one cannot exclude that the migration may have no effect at all on the migrant's health. The health conditions acquired in the place of origin could persist after migration. This is the *socialisation effect* whereby conditions and behaviours acquired at the place of origin, in particular during childhood, persist in later life whatever the new environment the migrant is exposed to (Kulu 2005). Adaptation effect may still exist but may not be sufficient to counterbalance the socialisation effect. The two effects, adaptation and socialisation, are therefore opposed.

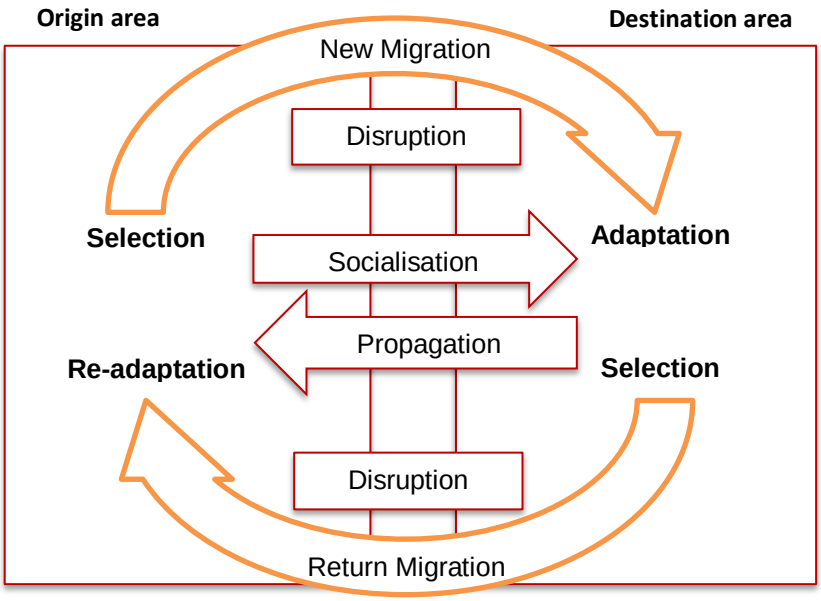
Research into the migration-health relationship often seeks to verify these four effects through empirical analysis. In the remainder of this paper we will refer to the SoSAD hypotheses to discuss the hypotheses that verify Socialisation, Selection, Adaptation and Disruption effects (Bocquier 2014). To note, the SoSAD hypotheses do not only apply to the study of diseases but have been extensively used since the 1960s to analyse reproductive health in relation to migration.

The SoSAD hypotheses have been associated with migration flows from less to more affluent areas, generally from rural to urban areas. These are the most common internal migration flows generally experienced by youth at the beginning of their working lives. Other flows have sometimes been considered and these have prompted the alternative "unhealthy return migrant" hypothesis, i.e. that of negative selection on health. For example, studies in South Africa have shown that prominently rural sending areas experience an excess mortality due to people 'returning home to die' (Clark et al. 2007; Collinson et al. 2009). The assumption is that new migrants will be attracted to places with better economic opportunities and living conditions, generally in urban areas, but that some migrants may partake in high risk behaviour (smoking, drinking, unhealthy diets, risky sexual encounters, violence) and may have difficulty accessing health services in these destinations, in particular if they reside in slums. This phenomenon is referred to as the segmented adaptation effect. As a consequence, the migrants will relocate to their place of origin when their health deteriorates to seek health care and support, thus contributing to higher mortality in rural areas. This return migration is also called the "midnight train" effect after a soul song that tells about a failed musician in Los Angeles who takes the midnight train back to Georgia, his place of origin (Nauman et al. 2015). This return of unhealthy migrants creates the so-called "salmon bias" that leaves the healthier at destination (Abraído-Lanza et al. 1999; Lu and Qin 2014).

Realising that the SoSAD hypotheses usually intended for new (first-time) migration may apply equally (but not necessarily with the same effect on health) to return migration, it is interesting to look at bi-directional migration flows between origin and destination areas. For simplicity, we present in Figure 1 the SoSAD hypotheses for two main migration flows, new migration and return migration. The selection and disruption hypotheses are generally synonymous whatever the direction of migration. However, the equivalent of socialisation, which refers mainly to behaviours and health conditions acquired during childhood, differs in the case of return migration, i.e. after the migrant spent some time at destination. We will refer to the *propagation effect* as the symmetrical effect to socialisation. This propagation (or diffusion) effect is conditional on adaptation to the place of destination and identifies the possibility that behaviour and health conditions at destinations can be spread to origin areas through return migration. After return migration, the migrant may re-adapt to its origin area, hence the *re-adaptation effect*.

There is an ambiguity with regards to the interplay between the environment and the behaviour of the migrants. For example, one may consider the exposure to a specific environment as fairly homogenous, while migrants may have different behaviours (segmented adaptation) that lead to negative selection by return migration or to positive selection through permanent settlement in the host area. The alternative explanation is that the environment is heterogeneous, exposing migrants to different risks depending on where they reside, while migrants may be uniformly positively selected through migration.

Figure 1: Interaction between migration and health before and after new migration or return migration



3. Theoretical Framework

The Table 1 presents for new (in-)migrants the different combinations of selection and differences in health risks exposure before and after migration, and the expected direction of the coefficients for two empirical variables controlling for length of exposure in the study area as compared to non-migrants (the reference category). These duration variables are easily constructed in longitudinal datasets. For return migrants, we can control for the exposure out of the site before return migration, in addition to short- and long-term exposure in the site after return migration. In statistical modelling terms, the selection and exposure effects form a 'data generating process': each combination of these effects generates a particular set of variable

coefficients. Table 1 presents every possible combination of these coefficients whose direction can be negative, positive or equal. These coefficients are taken to be measured all other non-migration-related variables being equal, i.e. they are measuring the net effect of varying exposure by migration status. The only hypothesis concerning mortality that we cannot test is that of disruption. Evidently, this is because we cannot measure mortality before and after migration for the same individuals. Therefore the Table 1 is does not include the disruption hypothesis depicted in Figure 1.

The respective effects of selection and exposure are easier to understand for return migrants. It is essential to note that "long exposure out" captures the propagation effect, i.e. the health risks brought about by exposure in the residence before return migration, while the exposure after return migration captures the difference in mortality between return migrants and non-migrants in the site. Taking the duration into account, the effect of "short exposure in" captures the selection effect, i.e. the difference between return migrants and non-migrants whatever the place of residence, since the health risks exposure before return migration is already controlled for. The re-adaptation effect is measured by "long exposure in", when mortality of return migrants gradually converges to that of non-migrants. The three effects of propagation, selection and re-adaptation are therefore measured net of one another and are displayed in Table 1 where the mortality of migrants is compared to that of non-migrants, the reference category.

For example, the column (1) assumes negative selection and no effect of long exposure out (i.e. no difference in health risks exposure before and after migration). It reads as follows: as compared to non-migrants in the study area, the migrants have a higher mortality within a given duration threshold (say, 2 years after their migration), i.e. "short exposure in" is ">", showing a negative selection effect through migration. It is followed by an adaptation effect ("long exposure in" is "=") meaning convergence towards non-migrants' mortality. Since there is no effect of "long exposure out", there is no propagation effect.

The column (15) reads as follows: the migrants have lower mortality in the first 2 years after their migration ("short exposure in" is "<"), i.e. positive selection effect since they reduced their health risks through migration. Mortality of the migrants remained the same thereafter (still "<"), there is no re-adaptation effect. Since "long exposure out" increased mortality (">"), the propagation effect is negative. The column (8) is for no selection effect but a positive propagation effect ("long exposure out" decreased mortality: "<"). If there is no selection, there cannot be re-adaptation effect and the "long exposure in" remains the same ("=").

This is a different story for new (in-)migrants since the exposure out is always long and its effect adds to the exposure in the destination site. Only the effect of short and long exposure after migration is available to identify the three hypotheses for in-migrants. Because all in-migrants are assumed to have had a long exposure in their place of origin, the interpretation of exposure in the site is conditional on knowledge on the difference of health risks before (B) and after (A) migration. This prior knowledge replaces the "long exposure out" effect observed for return migrants (the sign of the difference in health risks before and after migration for new migrants is the same as the effect of long exposure out for return migrants).

In other words, for in-migrants, the exposure in the site is confounded by exposure out of the site. The "short exposure in" effect is the sum of the selection effect and of the difference in health risks exposure before and after migration. The same is true for "long exposure in". In other words, the selection and adaptation effects are not measured net of the socialisation effect. If there is no difference in health risks in residences before and after migration (B=A), then the signs for short and long exposure in the site and their interpretation are the same as for return migrants (Table 1): there is no socialisation effect, the selection effect can be either negative (columns 1 and 2), nil (column 7), or positive (columns 10 and 11).

However, if health risks are lower before than after migration (B<A), then the signs for short and long exposure in the site will be completely different: the "<" sign will add to the "short" and "long exposure in" effects so that a negative selection will translate into a "=" sign (columns 3 and 4), no selection into a "<" sign (column 8), and positive selection into a "<<" sign (columns 12 and 13). For example, the column (8) assumes no selection effect but "short exposure in" is negative ("<") because of positive socialisation effect that decreases mortality; if there is no selection, there cannot be adaptation effect and the "long exposure in" remains the same ("<").

Conversely, if the health risks are higher before than after migration (B>A), the ">" sign will add to the "short" and "long exposure in" effects: a negative selection will translate into a ">>" sign (columns 5 and 6), no selection into a ">" sign (column 9), and positive selection into a "=" sign (columns 14 and 15). For example, the column (15) assumes positive selection but "short exposure in" is nil ("=") because of negative socialisation effect that increases mortality (B>A); migrants' mortality remained the same thereafter (still "="), there is no adaptation effect.

To note, for in-migrants, in absence of hypothesis on difference in health risks in residences before and after migration, the same empirical results may be generated by very different combinations of selection and exposure effects. For example, no short and long exposure effect in the site could mean negative selection, positive socialisation effect, and no adaptation (column 4), but could also mean no selection (hence no testable adaptation effect) and no socialisation (column 7), or positive selection, negative socialisation effect, and no adaptation (column 15). On the other hand, some empirical results can only be produced by one such combination. For example, very high "short exposure in" effect (">>") and "long exposure in" effect (">") can only mean negative selection, negative socialisation, and adaptation (column 5).

The identification of the right combination of selection and difference in exposure is easier for return migration than for in-migration thanks to the variable capturing long exposure out of the site. However, despite multiple potential interpretations, results for in-migrants may be less ambiguous than they appear with prior knowledge on the difference in health risks before (B) and after (A) migration. A reasonable assumption is that in rural areas new migrants are generally coming from neighbouring areas with the same level of health risks than the host area. In that case, we may reasonably assume no difference in health risks before and after migration (B=A), i.e. no socialisation effect. The combination of selection and adaptation hypotheses thus can easily be uniquely identified in column 1, 2, 7, 10, and 11. In urban areas, the assumption is much less valid because new migrants may come either from rural areas or from urban areas, knowing that urban areas are very heterogeneous. Consequently health risks may be assumed to be either higher before than after migration (B>A), equal (B=A), or lower (B<A).

The literature on new migration has largely assumed the combination of positive selection of migrants, negative socialisation effect, and adaptation (column 14). The literature on return migration refers to positive propagation effect, re-adaptation and negative selection in the "midnight train" case (column 3) or positive selection in the opposite case (column 12). However, there are potentially twelve other combinations when one considers the possibilities of no selection effect, no socialisation/propagation, and no (re-)adaptation. Therefore the observed patterns are expected to be much more diverse than found in the literature.

Table 1: Combinations of selection and exposure effects and observed mortality difference between migrants and non-migrants

Assumed selection	Negative selection (health before migration: migrant>non-migrant)						No selection (health before migration: migrant=non-migrant)			Positive selection (health before migration: migrant< non-migrant)					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Case of return migrants															
Observed difference in mortality risk after migration between return migrants and non-migrants:															
Short exposure in	>	>	>	>	>	>	=	=	=	<	<	<	<	<	<
Long exposure in	=	>	=	>	=	>	=	=	=	=	<	=	<	=	<
Re-adaptation effect (convergence with non-migrants)	yes	no	yes	no	yes	no	n.t.	n.t.	n.t.	yes	no	yes	no	yes	no
Observed long exposure out	=	=	<	<	>	>	=	<	>	=	=	<	<	>	>
Propagation effect	no	no	pos	pos	neg	neg	no	pos	neg	no	no	pos	pos	neg	neg
Case of new migrants															
Assumed difference in health risks exposure before (B) and after (A) migration	B=A	B=A	B<A	B<A	B>A	B>A	B=A	B<A	B>A	B=A	B=A	B<A	B<A	B>A	B>A
Observed difference in mortality risk after migration between new migrants and non-migrants:															
Short exposure in	>	>	=	=	>>	>>	=	<	>	<	<	<<	<<	=	=
Long exposure in	=	>	<	=	>	>>	=	<	>	=	<	<	<<	>	=
Inferred adaptation effect (convergence with non-migrants)	yes	no	yes	no	yes	no	n.t.	n.t.	n.t.	yes	no	yes	no	yes	no
Inferred socialisation effect (persistence of exposure B)	no	no	pos	pos	neg	neg	no	pos	neg	no	no	pos	pos	neg	neg

n.t.: not testable. pos: positive. neg: negative.

The Table 1 will be used to map the coefficients for the three exposure variables against the different combinations of these effects. In absence of direct information on the selection process and on health risks exposure before and after migration, these empirical variables will allow us to identify the most likely direction of the selection effect as well as the most likely change in health risks experienced by migrants before and after their move. This inference is conditioned on reasonable assumptions depending on the flow (new migration or return migration) and the site (rural or urban). Of course, it is also possible that migration-mortality relationship is not only generated by the combination of selection and exposure but also by other processes unaddressed, and therefore uncontrolled for in this paper. In addition, Table 1 only indicates the direction of the effects, not their magnitude. Any departure from the proposed patterns in Table 1 will be interpreted as a failure to explain the migration-mortality relationship with selection and exposure effects using the proposed theoretical framework.

4. Data and Methods

HDSSs have been developed in locations where population vital registration is poor or absent. The HDSS approach provides detailed, prospective, longitudinal data on demographic, health and socio-economic dynamics within these geographically-demarcated areas, usually the size of an administrative district. This is achieved by conducting a baseline census of the full population at the outset, with subsequent tracking of individual demographic events, namely, births, deaths, in- and out-migrations, on an on-going basis, at prescribed intervals within the study population. The HDSS platform enables a detailed exploration of determinants of two-way migration flows that can address this identified information gap.

The findings presented in this paper are based on data from nine HDSS sites that are members of the International Network for the Demographic Evaluation of Populations and Their Health (INDEPTH). The INDEPTH network is an initiative that has brought together HDSS sites from LMICs with the purpose of standardising data and techniques to enable cross-country comparative research (for more details concerning the methods and objectives of the INDEPTH organisation see Sankoh and Byass 2012). The HDSS sites represented in the current analysis are taken from four sub-Saharan African countries, and they represent a mix of rural, semi-rural and urban locations. The HDSS sites included in the study were selected based on data availability, and in order to maximise coverage of the Western, Eastern and Southern regions of the continent. These are Nanoro, Nouna and Ouagadougou in Burkina Faso; Kilifi, Kisumu and Nairobi in Kenya; Manhiça in Mozambique; Agincourt and Africa Centre in South Africa. These HDSS sites are part of the Multi-local Analysis of the Dynamics of Internal Migration And Health (MADIMAH) initiative within INDEPTH which commenced in 2011 with the aim of producing comparative studies on questions concerning migration and health (Gerritsen et al. 2013a). These sites represent one or more sub-district populations of their countries. Through detailed examinations of the dynamics of these demarcated geographical areas, insights about local-level migration flows may be elicited.

A profile of the nine HDSS sites is outlined in Table 2. The Nairobi HDSS is the most densely populated of the HDSS sites, with Nouna having the lowest population density. The Kilifi and Kisumu HDSS sites in rural Kenya have the largest populations under surveillance. The HDSS sites also differ in relation to the contiguity of the population under surveillance – the two urban sites in Nairobi and Ouagadougou comprise non-contiguous areas, while the rural-based sites are all contiguous. In order to provide some information on the context in which the sites are based, Table X (TO BE ADDED) presents a set of indicators of the four countries represented in the analyses.

For the purposes of this study, migration is defined as a move that crosses the geographical boundaries of the HDSS site (in either an inward or an outward direction). An in-migrant is an

individual who has lived in the HDSS area for at least 6 months, while an out-migrant is someone who moved away from the HDSS area for at least 6 months. Definitions of in- and out-migration may also differ according to HDSS site in relation to the specified time thresholds used to determine HDSS membership (varying from 3 to 6 months of residence within the boundaries of the HDSS). In order to achieve consistency, residency in an HDSS was standardised using the more conservative 6 month residency threshold. Consequently, migration as analysed in this study excludes moves within the HDSS boundaries, or moves for durations of less than 6 months outside of the HDSS area. For both first-time in-migrants and return migrants, exposure time following entry into the HDSS is categorised into durations of 6 - 24 months; 25 - 59 months and 60+ months. Return migrant exposure outside the HDSS discriminates between long exposure (taken at >36 months) and short exposure (<36 months).

The descriptive statistics and models were generated using Event History Analysis (EHA) techniques. This approach is appropriate for the examination of repeated events (such as migration) within the context of an individual's life course (Kulu and Milewski 2007; Yamaguchi 1991). Prior to commencing with EHA analyses, detailed data consistency and quality checks were conducted, and data were transformed into a biographical "residency episode" structure (Gerritsen et al. 2013a). This structure implies that events (including births, deaths, in- and out-migration and education measures) for individuals are recorded sequentially and in continuous-time (i.e. dates are attached to each event). The models allow for repeatable migration events over an individual's life course, thus individuals may contribute more than one migration event over time. In order to analyse in-migration, analysis time is reversed from age 75 (the upper age limit of the analysis) until the occurrence of an in-migration event, or to birth/enumeration if no in-migration event occurs (for further discussion of this method see Beguy, Bocquier and Zulu 2010). The analyses presented in this paper are based on data starting in 1998 or the earliest reliable year for migration analysis (see Table 3 for the different periods covered).

In- and out-migration rates were computed by 5-year age categories for each centre, stratified by sex. Rates are expressed as the number of events (in- or out-migrations) divided by the person time of the population at risk, expressed in years (person years at risk, or PYAR). In the case of out-migration, the population at risk corresponds to the time contributed by individuals within the HDSS over the corresponding age range. When an individual leaves and re-enters the population through return migration, the individual is included in the population at risk from the time of re-entry until censoring. For in-migration, this denominator represents the population at risk of "receiving" an in-migrant, and not the population from which the migrant originated.

Cox semi-parametric proportional hazards models were produced for each site in order to examine in-migration and return migration status as a determinant of death. These models control for age in the non- parametric part of the Cox model, and migration status, grouped calendar years and education as covariates. Models were stratified by sex to control for gender compositional effects. All analyses were performed using Stata version 13.

Table 2: HDSS sites included in this multi-centre analysis

HDSS Site	Population Size (approximate)	Size of Site (km²)	Settlement Type	Population Density Estimate (persons per km²)	Inception Year	Contiguity and Location
West Africa						
Nanoro HDSS Burkina Faso (Derra et al. 2012)	61 000	594.3	Rural	102.6	2009	Contiguous site situated in centre of Burkina Faso, 85km from capital, Ouagadougou
Nouna HDSS Burkina Faso (Sie et al. 2010)	84 336	1 756	(Mostly) Rural	48	1992	Contiguous site situated north west of Burkina Faso, 300km from capital, Ouagadougou
Ouagadougou HDSS Burkina Faso (Rossier et al. 2012)	81 717	14.73	Urban	5 547.7	2008	Non-contiguous site comprising three informal areas: Nonghin, Polesgo and Nioko 2, and two formal areas: Kilwin and Tanghin, north of city.
East Africa						
Kilifi HDSS Kenya (Scott et al. 2012)	261 919	900	(Mostly) Rural	291	2000	Contiguous site situated north of Mombasa on Indian Ocean coast of Kenya
Kisumu HDSS Kenya (Odhiambo et al. 2012)	223 406	700	(Mostly) Rural	319.2	2001	Contiguous site located in Rarieda, Siaya and Gem districts, northeast of Lake Victoria, Nyanza Province, western Kenya
Nairobi HDSS Kenya (Emina et al. 2011)	71 000	0.97	Urban	73 195.9	2002	Non-contiguous site comprising Viwandani and Korogocho slum settlements (7km apart) in capital, Nairobi
Southern Africa						
Africa Centre HDSS South Africa (Tanser et al. 2007)	85 000	438	Rural	194.1	1997	Contiguous site in the Umkanyakude district of KwaZulu-Natal
Agincourt HDSS South Africa (Kahn et al. 2012)	91 178	420	(Mostly) Rural	217.1	1992	Contiguous site situated in northeast South Africa close to border with Mozambique
Manhiça HDSS Mozambique (Sacarlal et al. 2009)	90 000	500	Rural	180	1996	Contiguous site located in southern Mozambique, 80 km north of capital, Maputo

Table 3 Characteristics of the analytical sample by HDSS site over the respective analysis periods

	Nouna HDSS		Ouagadougou HDSS Burkina Faso		Kilifi HDSS		Kisumu HDSS		Nairobi HDSS		Africa Centre HDSS		Agincourt HDSS	
	% Person Years		% Person Years		% Person Years		% Person Years		% Person Years		% Person Years		% Person Years	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Permanent Resident	157 687.21 62.58%	132 141.76 52.40%	58 798.02 83.51%	56 164.65 81.02%	315 957.67 56.63%	408 750.41 56.60%	294 541.55 74.77%	360 533.70 74.65%	86 538.36 43.87%	56 679.41 41.38%	116 444.01 64.93%	170 052.99 67.64%	81 680.05 45.17%	118 005.72 42.81%
In-Migrant														
6 - 24 months in HDSS	18 243.81 7.24%	29 353.31 11.64%	8 770.98 12.46%	9 941.43 14.34%	77 489.69 13.89%	95 949.36 13.29%	26 745.71 6.79%	39 179.08 8.11%	43 898.83 22.26%	32 375.38 23.64%	13 352.58 7.45%	16 773.85 6.67%	14 979.03 8.28%	30 646.21 11.12%
25 - 59 months in HDSS	23 245.42 9.22%	36 835.09 14.61%	2 842.88 4.04%	3 214.83 4.64%	77 761.70 13.94%	102 453.68 14.19%	27 286.80 6.93%	35 804.34 7.41%	34 295.33 17.39%	24 272.53 17.72%	15 059.87 8.40%	19 869.81 7.90%	21 283.35 11.77%	41 791.86 15.16%
60+ months in HDSS	28 573.94 11.34%	40 233.57 15.95%	~	~	51 413.48 9.22%	75 886.31 10.51%	12 554.56 3.19%	14 780.73 3.06%	12 217.18 6.19%	8 759.05 6.40%	13 081.05 7.29%	18 557.84 7.38%	41 421.31 22.91%	69 528.20 25.22%
Return Migrant														
6 - 24 months in HDSS	7 355.69 2.92%	4 710.15 1.87%	~	~	17 693.47 3.17%	18 773.46 2.60%	15 197.02 3.86%	15 732.88 3.26%	9 026.93 4.58%	6 665.32 4.87%	9 294.94 5.18%	10 657.67 4.24%	7 831.76 4.33%	5 545.94 2.01%
25 - 59 months in HDSS	9 232.28 3.66%	4 989.15 1.98%	~	~	13 918.96 2.49%	15 408.68 2.13%	13 691.54 3.48%	13 171.82 2.73%	8 347.64 4.23%	6 037.16 4.41%	8 457.28 4.72%	10 463.56 4.16%	7 266.36 4.02%	5 420.73 1.97%
60+ months in HDSS	7 646.77 3.03%	3 922.27 1.56%	~	~	3 665.84 0.66%	4 942.02 0.68%	3 902.31 0.99%	3 733.16 0.77%	2 921.76 1.48%	2 172.85 1.59%	3 643.20 2.03%	5 038.55 2.00%	6 356.28 3.52%	4 717.71 1.71%
Return Migrant Exposure >36months	246 006.28 97.63%	248 517.53 98.55%			548 149.07 98.25%	713 426.12 98.79%	388 637.16 98.66%	478 270.18 99.03%	194 946.66 98.83%	135 476.35 98.92%	174 293.86 97.19%	245 811.79 97.77%	168 513.30 93.19%	268 382.90 97.36%
36+ months away	5 978.84 2.37%	3 667.77 1.45%	~	~	9 751.74 1.75%	8 737.80 1.21%	5 282.32 1.34%	4 665.52 0.97%	2 299.37 1.17%	1 485.36 1.08%	5 039.06 2.81%	5 602.49 2.23%	12 304.83 6.81%	7 273.48 2.64%

	Nouna HDSS		Ouagadougou HDSS Burkina Faso		Kilifi HDSS		Kisumu HDSS		Nairobi HDSS		Africa Centre HDSS		Agincourt HDSS	
	% Person Years		% Person Years		% Person Years		% Person Years		% Person Years		% Person Years		% Person Years	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Period														
1 Jan 1998 - 1 Jan 2001 (1998)	29 518.17 11.71%	29 762.37 11.80%	~	~	~	~	~	~	~	~	~	~	39 432.52 21.81%	55 499.58 20.13%
1 Jan 2001 - 1 Jan 2004 (2001)	44 703.59 17.74%	44 526.59 17.66%	~	~	107 780.83 19.32%	137 178.64 19.00%	~	~	~	~	43 778.85 24.41%	61 698.05 24.54%	32 956.24 18.23%	52 441.51 19.02%
1 Jan 2004 - 1 Jan 2007 (2004)	54 232.43 21.52%	54 325.24 21.54%	~	~	135 271.40 24.25%	177 187.50 24.54%	114 93 065.52 23.63%	868.70 23.79%	40 60 030.76 30.43%	031.23 29.23%	43 210.36 24.10%	61 482.29 24.45%	32 712.73 18.09%	52 596.64 19.08%
1 Jan 2007 - 1 Jan 2010 (2007)	59 189.52 23.49%	59 471.15 23.58%	~	~	149 931.08 26.87%	195 820.99 27.12%	141 467.95 35.91%	172 354.44 35.69%	45 64 977.23 32.94%	184.60 32.99%	44 969.18 25.08%	63 770.97 25.36%	35 327.45 19.54%	55 600.20 20.17%
1 Jan 2010 - 1 Jan 2013 (2010)	64 341.42 25.53%	64 099.96 25.42%	70 411.87 100%	69 320.92 100%	164 917.51 29.56%	211 976.78 29.35%	159 386.01 40.46%	195 712.56 40.53%	51 72 238.03 36.62%	745.88 37.78%	47 374.54 26.42%	64 462.98 25.64%	40 389.20 22.34%	59 518.45 21.59%
Education														
No Formal - base	119 726.54 47.51%	149 753.05 59.38%	21 348.15 30.32%	29 178.50 42.09%	66 759.70 11.97%	292 929.32 40.56%	7 639.69 1.94%	52 594.88 10.89%	6 236.77 3.16%	9 468.25 6.91%	8 915.91 4.97%	19 976.46 7.95%	18 910.15 10.46%	55 734.17 20.22%
Some Primary	37 541.93 14.90%	20 998.05 8.33%	18 497.26 26.27%	14 719.60 21.23%	283 779.33 50.87%	254 246.23 35.21%	243 840.38 61.90%	299 926.39 62.10%	110 179.82 55.86%	86 218.15 62.95%	23 215.31 12.95%	47 039.13 18.71%	38 862.09 21.49%	50 950.15 18.48%
Some Secondary	18 733.95 7.43%	10 805.58 4.28%	20 252.98 28.76%	16 187.59 23.35%	65 220.03 11.69%	39 568.88 5.48%	89 834.49 22.81%	71 840.30 14.88%	39 76 576.83 38.82%	389.29 28.76%	98 220.10 54.77%	125 451.02 49.90%	113 128.91 62.57%	152 608.65 55.36%
Some Tertiary	852.12 0.34%	150.38 0.06%	3 777.84 5.37%	1 553.89 2.24%	12 910.62 2.31%	8 290.82 1.15%	16 186.47 4.11%	8 440.80 1.75%	2 732.71 1.39%	1 036.68 0.76%	46 617.47 25.99%	56 347.39 22.41%	8 317.20 4.60%	14 154.94 5.13%
Unknown	75 130.57 29.82%	70 478.24 27.95%	6 535.65 9.28%	7 681.34 11.08%	129 231.13 23.16%	127 128.67 17.60%	36 418.46 9.25%	50 133.33 10.38%	1 519.90 0.77%	849.34 0.62%	2 364.13 1.32%	2 600.28 1.03%	1 599.79 0.88%	2 208.47 0.80%

*** p<0.01, ** p<0.05, * p<0.1

The objectives of the comparison of nine HDSS sites are:

- To confirm the diversity of the migration-mortality relationship over a range of countries and type of residence across the continent;
- To confirm that the pattern of migration-mortality relationship is mainly generated by the combination of two processes: selection and exposure;
- To identify the most likely explanation for the patterns of mortality in local contexts characterised by high mobility and to check whether they conform to the well-known healthy migrant and unhealthy return migrant hypotheses;
- To help local health authorities to identify the categories of migrants for targeted interventions.

To note, the present comparative study does not aim at covering all existing situations on the continent, nor does it pretend to be representative. The situations are as exemplary as they can be considering the available data. Our hope is that these situations are diverse enough to inspire the analysis of other health issues in local areas where migration is important.

5. Results

5.1. Descriptive results

The rates of out-migration by site was analysed and reported on in a previous study (See Ginsburg et al. 2015). Across the same group of HDSS sites, between 7 and 21 per 100 PYAR of these local populations were found to have out-migrated between years 2009-2011. In-migration rates were of a similar magnitude with between 7 and 27 per 100 PYAR of individuals in-migrating over this period. However, both in- and out-migration rates were found to vary substantially by age group with the highest rates observed in early adult years (ages 15-29) for both males and females across all HDSS sites.

The probability to die between ages 15 and 60 (45q15) for males and females are presented for each HDSS site by period in Figures 3 and 4. In the 2010-2012 period for which data is available for all sites, Southern Africa sites experience the highest probability of male adult mortality over time, with probabilities between ages 15 and 60 being 0.66 and 0.65 for Agincourt and Manhica, closely followed by the Africa Centre at 0.60. Similarly, Southern African sites report the highest probabilities of mortality in females: 0.45 for Manhica, 0.39 for Agincourt and 0.37 for the Africa Centre. Lowest probabilities of mortality are evident within the Burkina Faso HDSS sites and Kilifi in East Africa. In all sites for which data is available for the years 2000 and more the probability is declining for both males and females.

Figure 2: Probability of Death between Ages 15 and 60 ($_{45}q_{15}$) by HDSS Site for Males

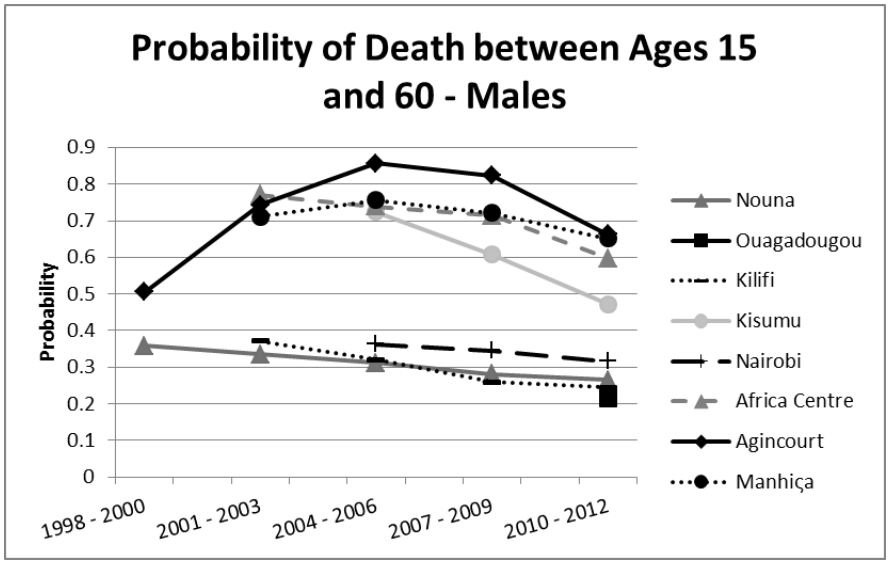
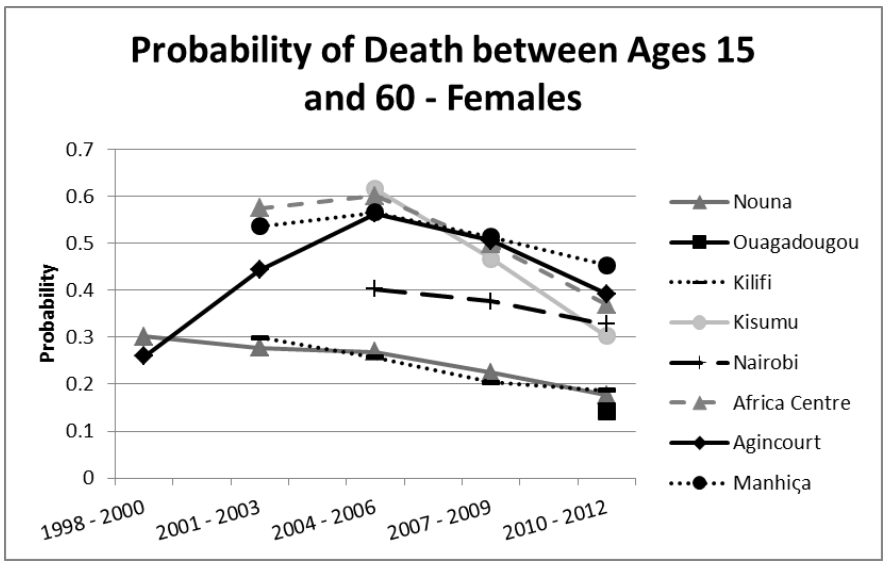


Figure 3: Probability of Death between Ages 15 and 60 ($_{45}q_{15}$) by HDSS Site for Females



5.2. Regression results

Rural Southern Africa

With regards to the Agincourt HDSS, no significant difference is observed between a first time entering in-migrant and a permanent resident. However, in the case of both male and female return migrants, the probability of death in the HDSS is 5.12 times higher for males and 5.51 times higher for females within two years following return, as compared to a permanent resident. Conversely in the Africa Centre HDSS, the effect of return migration is far less in magnitude for females (1.37 times higher between 6 and 24 months following return to the HDSS) and not apparent amongst males. The higher relative risk of mortality for female return migrants reduces the longer the duration spent in Africa Centre HDSS. Results from the Manhiça HDSS indicate that male and female first time in-migrants to the HDSS have a significantly higher risk of death within two years following their entry in the HDSS (1.28 times the risk for males and 1.59 times for females); however this risk reduces with length of stay in the HDSS. To note, risks within the first two years following return do not differ by gender for Agincourt and Manhiça HDSSs, but do differ for the Africa Centre.

The only case where a socialisation and propagation effects can be seen is amongst return migrants to the Agincourt HDSS. All other African HDSS show no evidence of these effects. The Agincourt HDSS situation probably reflects the propagation of the AIDS epidemic that affected this rural area particularly severely in the years 2000 (Bocquier et al. 2014). The propagation effect adds to the negative selection effect to result in particularly high mortality for return migrants as compared to non-migrants, which may be contrasted with the absence of migration effect for in-migrants. Here the "midnight train" or negative selection effect (unhealthy return migrant) is adding to a propagation effect whereby migrants are bringing home higher health risks acquired outside the study area.

Still in South Africa, the Africa Centre HDSS shows no migration effect whatsoever for males. This means that, among males, in-migrants as well as return migrants are moving from areas where they faced similar exposure than non-migrants in the site and that there is no selection effect. Female return migrants, however, show a negative selection effect. Female in-migrants in Africa Centre HDSS are the only group amongst all African HDSS under study that present a pattern that does not conform to any combination displayed in Table 1. This group experiences lower mortality for short exposure in the site but higher mortality after 5 years of residence: this reversal of trends after long exposure in the site is not predicted by our hypotheses and would therefore need further investigation.

In the Manhiça HDSS in Mozambique, the dominant pattern is that of negative selection with (re-)adaptation, and no socialisation/propagation effect. The "unhealthy migrant" hypothesis applies in Manhiça to both in-migrants and return migrants. The number of female return migrants was too few to compute meaningful regression analysis coefficients.

Rural East Africa

The results from the Kisumu HDSS reveal both first time in-migrants and return migrants have a higher risk of mortality within the first two years following entry into the HDSS as compared with non-migrants, with the risk for females being higher than for males (for first-time in-migrants 1.80 times the risk and 1.35 times the risk respectively). The risk of mortality declines with duration of residence in the HDSS. Conversely, in-migrants and return migrants to the Kilifi HDSS are positively selected on health with lower risk of mortality within two years following entry to the HDSS (male return migrants have 0.56 and females 0.62 times the risk of death during this period).

The gender difference in mortality pattern is negligible in these two sites. The difference by migration status is also absent: the combination of effects is the same for in-migrants and return migrants. Both sites show (re-)adaptation effect but no socialization/propagation effect, meaning that migrants faced similar health risks where they migrated from than the non-migrants in the site. However Kisumu and Kilifi HDSS differ markedly in terms of the selection hypothesis: it is positive in the case of Kilifi (conforming to the "healthy migrant" hypothesis) and negative in the case of Kisumu (conforming to the "unhealthy migrant" hypothesis). The situation in Kisumu HDSS in Kenya is very similar to that of Manhiça HDSS in Mozambique, while the Kilifi HDSS situation is unique among the African sites under study.

Rural West Africa

In Burkina Faso, males returning to the Nouna HDSS, or entering the Nouna HDSS for the first time are positively selected on health with their risk of mortality being 0.55 and 0.72 times the risk of non-migrants respectively. For females, no significant relationship between mortality and migrant status is observed.

There is a sharp contrast between males and females in Nouna HDSS: females show no migration effect, whereas males show a positive selection effect that persists over time, i.e. with no (re-)adaptation effect.

ADD NANORO

Urban Sites: Ouagadougou and Nairobi

With respect to the Ouagadougou HDSS male first time entering migrants have 0.61 times the risk of death within two years following entry to the HDSS as compared to non-migrants. The risk converges to that of non-migrants after some years. This pattern is only compatible with no difference in health risks before and after migration, i.e. with positive selection, adaptation and no socialisation effect. This is also the case for male in-migrants in Nairobi HDSS (relative risks of 0.77 times within two years of entry as compared to non-migrants).

The number of return migrants is too low in the Ouagadougou HDSS due to the site's more recent inception date. Males returning to the Nairobi HDSS show the opposite risks to in-migrants: their risk of death is 1.33 times higher than non-migrants in the HDSS, while the risk converges to that of non-migrants thereafter. This pattern is compatible with negative selection, re-adaptation and no propagation effect.

To note, for both males in Ouagadougou and Nairobi, the observed patterns are only compatible with no difference in health risks before and after migration. In other words, migrants faced similar health risks where they migrated from than the non-migrants in these two sites. This is also the case of females in Nairobi but there subsists a doubt about females in Ouagadougou: the absence of migration effect for female in-migrants is compatible with all situations of difference in health risks before and after migration. However, there is little reason to believe that female migrants were subjected to very different health risks than male migrants. Therefore, assuming no socialisation, female in-migrants pattern is compatible with no selection effect, and therefore not testable adaptation effect.

Contrary to males, female in-migrants in Nairobi HDSS show a negative selection effect, the risk of mortality within the first two years is 1.25 times the risk for a permanent resident, but their risk converge to that of non-migrants thereafter (adaptation effect).

Table 4 Cox proportional hazards models – West African HDSSs

	Nouna HDSS		Ouagadougou HDSS Burkina Faso	
	All Deaths		All Deaths	
	Male	Female	Male	Female
Permanent Resident				
In-Migrant				
6 - 24 months in HDSS	0.72***	0.90	0.61**	0.93
	(0.58 - 0.90)	(0.75 - 1.08)	(0.38 - 0.97)	(0.59 - 1.47)
25 - 59 months in HDSS	0.82**	0.95	1.12	0.96
	(0.69 - 0.98)	(0.81 - 1.11)	(0.62 - 2.00)	(0.45 - 2.04)
60+ months in HDSS	0.77***	0.85**	~	~
	(0.66 - 0.91)	(0.73 - 1.00)		
Return Migrant				
6 - 24 months in HDSS	0.55***	0.83	~	~
	(0.37 - 0.82)	(0.58 - 1.19)		
25 - 59 months in HDSS	0.50***	0.53***	~	~
	(0.36 - 0.71)	(0.36 - 0.79)		
60+ months in HDSS	0.39***	0.55***	~	~
	(0.28 - 0.55)	(0.38 - 0.82)		
Return Migrant Exposure >36months				
36+ months away	1.34	1.33	~	~
	(0.90 - 2.00)	(0.87 - 2.02)		
Period				
1 Jan 1998 - 1 Jan 2001 (1998)	0.77***	0.98	~	~
	(0.67 - 0.90)	(0.83 - 1.15)		
1 Jan 2001 - 1 Jan 2004 (2001)	0.78***	1.06	~	~
	(0.68 - 0.90)	(0.92 - 1.23)		
1 Jan 2004 - 1 Jan 2007 (2004)	0.88**	1.14*	~	~
	(0.77 - 1.00)	(0.99 - 1.32)		
1 Jan 2007 - 1 Jan 2010 (2007)	1.01	1.11	~	~
	(0.89 - 1.15)	(0.96 - 1.28)		
1 Jan 2010 - 1 Jan 2013 (2010)	1.00	1.00	1	1
	(1.00 - 1.00)	(1.00 - 1.00)	(1.00 - 1.00)	(1.00 - 1.00)
Education				
No Formal - base	~	~	~	~
	~	~	~	~
Some Primary	1.36***	1.22	1.13	1.13
	(1.15 - 1.62)	(0.94 - 1.60)	(0.84 - 1.51)	(0.75 - 1.69)
Some Secondary	1.17	0.79	1.14	0.7
	(0.87 - 1.57)	(0.47 - 1.32)	(0.82 - 1.58)	(0.41 - 1.19)
Some Tertiary	0.62	0.00	0.96	0.36
	(0.15 - 2.49)	(0.00 - 0.00)	(0.52 - 1.80)	(0.05 - 2.60)
Unknown	4.78***	4.62***	1.43	1.59*
	(4.32 - 5.27)	(4.19 - 5.11)	(0.93 - 2.20)	(0.99 - 2.54)
Observations	369 512	383 136	40,696	40,882
Wald Chi-square	1063	1064	7.587	8.546
Log Lik	-16122	-14956	-2050	-1268
Subjects	45864	51906	33377	34174
Time at risk	251985	252185	70412	69321
Failures	2130	1948	317	195

*** p<0.01, ** p<0.05, * p<0.1

Table 5 Cox proportional hazards models – East African HDSSs

	Kilifi HDSS		Kisumu HDSS		Nairobi HDSS	
	All Deaths		All Deaths		All Deaths	
	Male	Female	Male	Female	Male	Female
Permanent Resident						
In-Migrant						
6 - 24 months in HDSS	0.59***	0.58***	1.35***	1.80***	0.77***	1.25**
	(0.53 - 0.66)	(0.52 - 0.64)	(1.23 - 1.49)	(1.64 - 1.97)	(0.66 - 0.90)	(1.05 - 1.50)
25 - 59 months in HDSS	0.67***	0.74***	1.12**	1.26***	0.91	1.01
	(0.60 - 0.74)	(0.67 - 0.82)	(1.01 - 1.24)	(1.13 - 1.40)	(0.78 - 1.07)	(0.83 - 1.24)
60+ months in HDSS	0.94	1.11*	0.95	1.17*	0.83	0.90
	(0.84 - 1.06)	(0.99 - 1.24)	(0.81 - 1.12)	(0.99 - 1.39)	(0.65 - 1.05)	(0.67 - 1.22)
Return Migrant						
6 - 24 months in HDSS	0.56***	0.62***	1.36***	1.53***	1.31**	1.16
	(0.44 - 0.70)	(0.49 - 0.77)	(1.17 - 1.57)	(1.31 - 1.79)	(1.04 - 1.65)	(0.85 - 1.57)
25 - 59 months in HDSS	0.76**	0.86	1.40***	1.13	1.10	0.97
	(0.60 - 0.96)	(0.68 - 1.07)	(1.21 - 1.62)	(0.94 - 1.36)	(0.86 - 1.41)	(0.70 - 1.35)
60+ months in HDSS	0.74	0.92	1.15	1.12	1.33	1.39
	(0.49 - 1.12)	(0.63 - 1.33)	(0.87 - 1.52)	(0.79 - 1.60)	(0.93 - 1.89)	(0.91 - 2.14)
Return Migrant Exposure >36months						
36+ months away	0.95	1.11	1.04	0.88	1.46*	1.36
	(0.71 - 1.29)	(0.84 - 1.48)	(0.81 - 1.33)	(0.64 - 1.23)	(0.99 - 2.15)	(0.77 - 2.40)
Period						
1 Jan 1998 - 1 Jan 2001 (1998)	~	~	~	~	~	~
	~	~	~	~	~	~
1 Jan 2001 - 1 Jan 2004 (2001)	0.96	1.06	~	~	~	~
	(0.87 - 1.06)	(0.96 - 1.17)	~	~	~	~
1 Jan 2004 - 1 Jan 2007 (2004)	1.16***	1.31***	1.80***	2.21***	1.05	1.33***
	(1.06 - 1.27)	(1.20 - 1.42)	(1.68 - 1.92)	(2.07 - 2.36)	(0.92 - 1.20)	(1.13 - 1.57)
1 Jan 2007 - 1 Jan 2010 (2007)	0.99	0.98	1.39***	1.57***	1.07	1.03
	(0.91 - 1.09)	(0.90 - 1.06)	(1.31 - 1.48)	(1.48 - 1.68)	(0.95 - 1.21)	(0.88 - 1.20)
1 Jan 2010 - 1 Jan 2013 (2010)	1.00	1.00	1.00	1.00	1.00	1.00
	(1.00 - 1.00)	(1.00 - 1.00)	(1.00 - 1.00)	(1.00 - 1.00)	(1.00 - 1.00)	(1.00 - 1.00)
Education						
No Formal - base	~	~	~	~	~	~
	~	~	~	~	~	~
Some Primary	0.89**	1.05	0.70***	0.81***	0.95	0.90
	(0.81 - 0.99)	(0.94 - 1.17)	(0.63 - 0.78)	(0.75 - 0.87)	(0.76 - 1.19)	(0.73 - 1.11)
Some Secondary	0.83**	0.91	0.53***	0.55***	0.62***	0.55***
	(0.71 - 0.96)	(0.73 - 1.14)	(0.47 - 0.60)	(0.49 - 0.62)	(0.49 - 0.79)	(0.43 - 0.70)
Some Tertiary	0.72**	0.66*	0.36***	0.32***	0.60*	0.15*
	(0.55 - 0.93)	(0.42 - 1.06)	(0.30 - 0.43)	(0.24 - 0.43)	(0.33 - 1.08)	(0.02 - 1.07)
Unknown	4.10***	8.39***	0.75***	0.92*	1.70**	1.70*
	(3.77 - 4.47)	(7.80 - 9.01)	(0.66 - 0.85)	(0.84 - 1.01)	(1.07 - 2.68)	(0.95 - 3.03)
Observations	2 382 427	3 097 890	292 304	352 742	370 927	266 241
Wald Chi-square	2034	3974	623.3	969	105	82.61
Log Lik	-34688	-39304	-53320	-56146	-12215	-7698
Subjects	145669	168010	98838	123054	67859	50049
Time at risk	557901	722164	393920	482936	197246	136962
Failures	4145	4548	6345	6396	1511	992

*** p<0.01, ** p<0.05, * p<0.1

Table 6 Cox proportional hazards models – Southern African HDSSs

	Africa Centre HDSS		Agincourt HDSS		Manhiça HDSS	
	All Deaths		All Deaths		All Deaths	
	Male	Female	Male	Female	Male	Female
Permanent Resident						
In-Migrant						
6 - 24 months in HDSS	0.93	0.86**	0.91	1.04	1.42***	1.45***
	(0.83 - 1.05)	(0.76 - 0.97)	(0.78 - 1.08)	(0.91 - 1.20)	(1.27 - 1.59)	(1.27 - 1.65)
25 - 59 months in HDSS	0.86**	1.10	1.00	1.06	1.13**	1.19***
	(0.76 - 0.97)	(0.98 - 1.24)	(0.86 - 1.15)	(0.94 - 1.20)	(1.02 - 1.26)	(1.06 - 1.35)
60+ months in HDSS	1.00	1.23***	0.92	1.01	1.01	1.08
	(0.87 - 1.15)	(1.07 - 1.41)	(0.82 - 1.05)	(0.92 - 1.11)	(0.90 - 1.12)	(0.96 - 1.21)
Return Migrant						
6 - 24 months in HDSS	1.18*	1.37***	5.12***	5.51***	1.28***	1.59***
	(1.00 - 1.40)	(1.15 - 1.63)	(4.49 - 5.83)	(4.72 - 6.44)	(1.11 - 1.49)	(1.37 - 1.85)
25 - 59 months in HDSS	1.08	1.25**	1.34***	1.77***	1	1.47***
	(0.91 - 1.29)	(1.04 - 1.50)	(1.14 - 1.57)	(1.44 - 2.17)	(0.87 - 1.17)	(1.28 - 1.69)
60+ months in HDSS	1.03	1.27*	1.07	0.83	1.1	1.08
	(0.80 - 1.33)	(0.99 - 1.64)	(0.91 - 1.26)	(0.63 - 1.09)	(0.93 - 1.29)	(0.90 - 1.28)
Return Migrant Exposure >36months						
36+ months away	0.97	1.14	1.41***	1.34***	1.16*	1.16
	(0.77 - 1.23)	(0.89 - 1.44)	(1.26 - 1.59)	(1.14 - 1.58)	(0.97 - 1.39)	(0.96 - 1.41)
Period						
1 Jan 1998 - 1 Jan 2001 (1998)	~	~	0.85**	0.76***	~	~
	~	~	(0.74 - 0.97)	(0.67 - 0.87)		
1 Jan 2001 - 1 Jan 2004 (2001)	1.06	1.37***	1.52***	1.23***	1.13**	1.30***
	(0.95 - 1.18)	(1.23 - 1.52)	(1.35 - 1.71)	(1.09 - 1.38)	(1.02 - 1.24)	(1.18 - 1.43)
1 Jan 2004 - 1 Jan 2007 (2004)	1.31***	1.77***	1.98***	1.57***	1.25***	1.39***
	(1.18 - 1.45)	(1.60 - 1.96)	(1.78 - 2.21)	(1.41 - 1.75)	(1.14 - 1.36)	(1.28 - 1.52)
1 Jan 2007 - 1 Jan 2010 (2007)	1.25***	1.42***	1.75***	1.31***	1.13***	1.21***
	(1.13 - 1.39)	(1.29 - 1.57)	(1.57 - 1.94)	(1.18 - 1.46)	(1.04 - 1.24)	(1.11 - 1.31)
1 Jan 2010 - 1 Jan 2013 (2010)	1.00	1.00	1.00	1.00	1	1
	(1.00 - 1.00)	(1.00 - 1.00)	(1.00 - 1.00)	(1.00 - 1.00)	(1.00 - 1.00)	(1.00 - 1.00)
Education						
No Formal - base	~	~	~	~	~	~
	~	~	~	~		
Some Primary	1.19***	1.03	0.95	0.89**	0.83***	1.02
	(1.06 - 1.34)	(0.92 - 1.15)	(0.86 - 1.05)	(0.81 - 0.99)	(0.77 - 0.90)	(0.95 - 1.10)
Some Secondary	0.85**	0.88**	0.84***	0.88**	0.59***	0.53***
	(0.75 - 0.96)	(0.78 - 0.99)	(0.76 - 0.94)	(0.79 - 0.98)	(0.51 - 0.67)	(0.43 - 0.65)
Some Tertiary	0.36***	0.30***	0.41***	0.49***	~	~
	(0.31 - 0.41)	(0.26 - 0.35)	(0.34 - 0.51)	(0.40 - 0.60)		
Unknown	7.02***	10.60***	3.80***	3.23***	0	0
	(6.12 - 8.05)	(9.26 - 12.13)	(3.07 - 4.72)	(2.52 - 4.14)	(0.00 - 0.00)	(0.00 - 0.00)
Observations	132 397	178 581	334 011	486 753	161,391	254,925
Wald Chi-square	1792	2241	2296	1122	195.2	236.2
Log Lik	-26421	-30509	-22795	-25388	-30180	-34678
Subjects	38234	46303	40818	55904	37663	48787
Time at risk	179333	251414	180818	275656	171254	270686
Failures	3593	3860	3198	3143	3966	4203

*** p<0.01, ** p<0.05, * p<0.1

6. Discussion and conclusion

The results confirm the diversity of the migration-mortality relationship over a range of rural and urban local areas in three African regions. The selection and exposure effects are very diverse

across the continent and within each country. No single pattern fits all situations: only two sites (Manhiça and Kisumu HDSS) have similar situations although being very distant. Gender differences are absent in about half the sites.

The results also confirm that the pattern of migration-mortality relationship is mainly generated by the combination of two processes: selection and exposure (in and out of the site). Out of 30 observed patterns there is only one (female in-migrants in Africa Centre HDSS) that does not conform to the expected combination of selection and exposure. Therefore, the proposed theoretical framework proves valid, and quite effective in interpreting the data at hand. However, an important limitation is that for in-migrants into rural areas, prior knowledge on origin areas (assumed difference in health risks before and after first in-migration) is necessary for interpreting the data.

As regard to the selection hypothesis, the results present a range of situations. The healthy in-migrant hypothesis is confirmed in four sites out of eight for males (Kilifi, Nouna, Ouagadougou and Nairobi HDSSs) and in only one site for females (Kilifi HDSS). It is contradicted in two sites (Manhiça and Kisumu HDSSs, for both males and females) while migration has no effect for in-migrants in two sites for males (Agincourt and Africa Centre HDSSs) and two sites for females (Agincourt and Nouna HDSSs).

The pattern amongst return migrants corresponds exactly to that of in-migrants in all sites but Africa Centre HDSS for females (which pattern is not consistent with our theoretical framework) as well as in Nairobi HDSS, where male return migrants show negative selection (as opposed to positive selection for in-migrants) and female return migrants show no selection (as opposed to negative selection for in-migrants). The similar patterns for in-migrants and return migrants found in most sites show that the nature of the migration does not markedly influence the interplay between selection and exposure effects. In particular, the "healthy migrant" does not oppose the "unhealthy return migrant" in the same site, except for males in one urban site, Nairobi HDSS.

How might these results impact on the administration of public health services in these HDSS sites, in particular as relating to migrants? Two effects call for particular attention: negative selection and negative socialisation/propagation. Migrants who are negatively selected on health are clearly a concern in Manhiça and Kisumu, whatever the migration status (in- or return migrant) and gender. In the Agincourt HDSS, male and female return migrants should be also targeted, especially since they are also vectors of negative propagation in the site. In the Africa Centre HDSS, female return migrants' health is a concern. In Nairobi HDSS, the concern is for male return migrants and for female in-migrants. In all these cases, the migrants should be targeted in the first 2 or 3 years after their arrival to the HDSS. After some years, migrants' risks tend to converge with those of non-migrants. The public health intervention would then help to reduce the risks upon arrival and accelerate the convergence.

Conditional on the validity of our assumption of no difference in health risks before and after first in-migration, the negative socialisation/propagation effect is not a concern in any of the sites under study, except in Agincourt HDSS for return migrants. There, in absence of means to reduce health risks in places of destination of migrants, prevention targeted towards residents who intend to migrate could help reduce health risks taken in migration destinations. Contrary to the expected, return migrants to rural areas (presumably returning from more affluent areas, in cities or abroad) do not appear to be vectors of positive propagation, as this effect is not evident in these data. Seeking support of return migrants with long exposure out of the site to implement local public health policy would not help so much.

Table 7: Selection, (re-)adaptation, propagation/socialisation hypotheses and actual mortality difference between non-migrant and migrants males depending on direction of migration in rural HDSS

Sites	Agincourt	Africa C	Manhiça	Kilifi	Kisumu	Nouna	Nanoro
New in-migrants							
Short exposure in	=	=	>	<	>	<	
Long exposure in	=	=	=	=	=	<	
Compatible hypotheses (assuming No social.)	No sel. Adap. n.t. No social.	No sel. Adap. n.t. No social.	Neg sel. Adap. No social.	Pos sel. Adap. No social.	Neg sel. Adap. No social.	Pos sel. No adap. No social.	
Return migrants							
Short exposure in	>	=	>	<	>	<	
Long exposure in	=	=	=	=	=	<	
Long exposure out	>	=	=	=	=	=	
Compatible hypotheses	Neg sel. Re-adap. Neg prop.	No sel. Re-adap. n.t. No prop.	Neg sel. Re-adap. No prop.	Pos sel. Re-adap. No prop.	Neg sel. Re-adap. No prop.	Pos sel. No re-adap. No prop.	

Social.: socialisation; prop.: propagation; social.: socialisation; sel.: selection; adap.: adaptation; n.t.: not testable ; n.a.: not available

Table 8: Selection, (re-)adaptation, propagation/socialisation hypotheses and actual mortality difference between non-migrant and migrants males depending on direction of migration and assumed difference in health risks before (B) and after (A) migration in urban HDSS

Sites	Ouagadougou			Nairobi		
New in-migrants						
Assumed difference in health risks exposure before (B) and after (A) migration	B=A	B<A	B>A	B=A	B<A	B>A
Short exposure in	<	<	<	<	<	<
Long exposure in	=	=	=	=	=	=
Compatible hypotheses	Pos sel. Adap. No social.	None	None	Pos sel. Adap. No social.	None	None
Return migrants						
Short exposure in		n.a.		>	>	>
Long exposure in		n.a.		=	=	=
Long exposure out		n.a.		=	=	=
Compatible hypotheses		n.a.		Neg sel. Re-adap. No prop.	None	None

Social.: socialisation; prop.: propagation; social.: socialisation; sel.: selection; adap.: adaptation; n.t.: not testable ; n.a.: not available

Table 9: Selection, (re-)adaptation, propagation/socialisation hypotheses and actual mortality difference between non-migrant and migrants females depending on direction of migration in rural HDSS

Sites	Agincourt	Africa C	Manhiça	Kilifi	Kisumu	Nouna	Nanoro
New in-migrants							
Short exposure in	=	<	>	<	>	=	
Long exposure in	=	>	=	=	=	=	
Compatible hypotheses (assuming No social.)	No sel. Adapt. n.t. No social.	None	Neg. sel. Adapt. No social.	Pos. sel. Adapt. No social.	Neg. sel. Adapt. No social.	No sel. Adapt. n.t. No social.	
Return migrants							
Short exposure in	>	>	n.a.	<	>	=	
Long exposure in	=	=	n.a.	=	=	=	
Long exposure out	>	=	n.a.	=	=	=	
Compatible hypotheses	Neg. sel. Re-adapt. Neg prop.	Neg. sel. Re-adapt. No prop.	n.a.	Pos. sel. Re-adapt. No prop.	Neg. sel. Re-adapt. No prop.	No sel. Adapt. n.t. No prop.	

Social.: socialisation; prop.: propagation; social.: socialisation; sel.: selection; adap.: adaptation; n.t.: not testable ; n.a.: not available

Table 10: Selection, (re-)adaptation, propagation/socialisation hypotheses and actual mortality difference between non-migrant and migrants females depending on direction of migration and assumed difference in health risks before (B) and after (A) migration in urban HDSS

Sites	**Ouagadougou**			**Nairobi**		
New in-migrants						
Assumed difference in health risks before (B) and after (A) migration	B=A	B<A	B>A	B=A	B<A	B>A
Short exposure in	=	=	=	>	>	>
Long exposure in	=	=	=	=	=	=
Compatible hypotheses	No sel. Adapt. n.t. No social.	Neg. sel. No adapt. Pos social.	Pos. sel. No adapt. Neg social.	Neg. sel. Adapt. No social.	None	None
Return migrants						
Short exposure in		n.a.		=	=	=
Long exposure in		n.a.		=	=	=
Long exposure out		n.a.		=	=	=
Compatible hypotheses		n.a.		No sel. Re-adapt. n.t. No prop.	None	None

Social.: socialisation; prop.: propagation; social.: socialisation; sel.: selection; adap.: adaptation; n.t.: not testable ; n.a.: not available

7. References

Abraído-Lanza, A.F., B.P. Dohrenwend, D.S. Ng-Mak, and J.B. Turner. 1999. "The Latino mortality paradox: a test of the "salmon bias" and healthy migrant hypotheses." *American Journal of Public Health* 89(10):1543-1548.

Beguy, D., P. Bocquier, and E.M. Zulu. 2010. "Circular migration patterns and determinants in Nairobi slum settlements." *Demographic Research* 23(20):549-586.

Bell, M., E. Charles-Edwards, P. Ueffing, J. Stillwell, M. Kupiszewski, and D. Kupiszewska. 2015. "Internal Migration and Development: Comparing Migration Intensities Around the World." *Population and Development Review* 41(1):33-58.

Bocquier, P. 2014. "Longitudinal surveillance challenges and consequences of migration for health of individuals and communities." in *2014 Columbia Global Migration Network Conference Schedule*. New York.

Bocquier, P., M. Collinson, S.J. Clark, A. Gerritsen, K. Kahn, and S.M. Tollman. 2014. "Ubiquitous burden: the contribution of migration to AIDS and Tuberculosis mortality in rural South Africa." *African Population Studies* 28(1):691-701.

Boutayeb, A. 2006. "The double burden of communicable and non-communicable diseases in developing countries." *Transactions of The Royal Society of Tropical Medicine and Hygiene* 100(3):191-199.

Choi, K. 2014. "Fertility in the context of Mexican migration to the United States." *Demographic Research* 30:703-738.

Clark, S., M. Collinson, K. Kahn, K. Drullinger, and S. Tollman. 2007. "Returning home to die: circular labour migration and mortality in South Africa." *Scandinavian Journal of Public Health* 35 (Supplement 69):35-44.

Collinson, M., K. Adazu, M. White, and S.E. Findley. 2009. "The Dynamics of Migration, Health and Livelihoods: INDEPTH Network perspectives." Farnham: Ashgate.

Collinson, M.A., M.J. White, P. Bocquier, S.J. McGarvey, S.A. Afolabi, S.J. Clark, K. Kahn, and S.M. Tollman. 2014. "Migration and the Epidemiological Transition: Insights from the Agincourt sub-district of northeast South Africa." *Global Health Action* 7(23514):1-15.

Ezzati, M., S. Vander Hoorn, C.M.M. Lawes, R. Leach, W.P.T. James, A.D. Lopez, A. Rodgers, and C.J.L. Murray. 2005. "Rethinking the "Diseases of Affluence" Paradigm: Global Patterns of Nutritional Risks in Relation to Economic Development." *PLoS Med* 2(5):e133.

Gerritsen, A., P. Bocquier, M. White, C. Mbacke, N. Alam, D. Beguy, F. Odhiambo, C. Sacoor, H.D. Phuc, S. Punpuing, and M.A. Collinson. 2013a. "Health and Demographic Surveillance Systems: Contributing to an understanding of the dynamics in migration and health." *Glob Health Action* 6:21496.

Gerritsen, A., P. Bocquier, M. White, C. Mbacké, N. Alam, D. Beguy, F. Odhiambo, C. Sacoor, H.D. Phuc, S. Punpuing, and M.A. Collinson. 2013b. "Health and demographic surveillance systems: contributing to an understanding of the dynamics in migration and health." *Global Health Action* 6(21496):1-8.

Goldstein, S. 1973. "Interrelation Between Migration and Fertility in Thailand." *Demography* 10(2):225-241.

Harpham, T. 2009. "Urban health in developing countries: What do we know and where do we go?" *Health & Place* 15(1):107-116.

Hervitz, H.M. 1985. "Selectivity, Adaptation, or Disruption? A Comparison of Alternative Hypotheses on the Effects of Migration on Fertility: The Case of Brazil." *International Migration Review* 19(2):293-317.

Kulu, H. 2005. "Migration and Fertility: Competing Hypotheses Re-examined." *European Journal of Population / Revue Européenne de Démographie* 21(1):51-87.

Kulu, H.and N. Milewski. 2007. "Family change and migration in the life course: an introduction." *Demographic Research* 17(19):567-590.

Kulu, H.and F. Steele. 2013. "Interrelationships between childbearing and housing transitions in the family life course." *Demography* 50(5):1687-1714.
Lu, Y.and L. Qin. 2014. "Healthy migrant and salmon bias hypotheses: A study of health and internal migration in China." *Social Science & Medicine* 102:41-48.
Nauman, E., M. VanLandingham, P. Anglewicz, U. Patthavanit, and S. Punpuing. 2015. "Rural-to-urban migration and changes in health among young adults in Thailand." *Demography* 52(1):233-257.
Salomon, J.A.and C.J.L. Murray. 2002. "The Epidemiologic Transition Revisited: Compositional Models for Causes of Death by Age and Sex." *Population and Development Review* 28(2):205-228.
Sankoh, O.and P. Byass. 2012. "The INDEPTH Network: Filling vital gaps in global epidemiology." *International Journal of Epidemiology* 41(3):579-588.
United Nations Population Division. 2014. "World Population Prospects: The 2012 Revision."
Urquia, M.L.and A.J. Gagnon. 2011. "Glossary: migration and health." *Journal of Epidemiology and Community Health*:jech. 2010.109405.
White, M.and D. Lindstrom. 2006. "Internal migration. InD. L. Poston & M. Micklin (Eds.), Handbook of population (pp. 331-346)." New York: Springer.
Yamaguchi, K. 1991. *Event history analysis*. Newbury Park: Sage Publications.
Yusuf, S., S. Reddy, S. Ôunpuu, and S. Anand. 2001. "Global Burden of Cardiovascular Diseases: Part I: General Considerations, the Epidemiologic Transition, Risk Factors, and Impact of Urbanization." *Circulation* 104(22):2746-2753.