



Demographic Research a free, expedited, online journal of peer-reviewed research and commentary in the population sciences published by the Max Planck Institute for Demographic Research Konrad-Zuse Str. 1, D-18057 Rostock · GERMANY www.demographic-research.org

DEMOGRAPHIC RESEARCH

**VOLUME 16, ARTICLE 8, PAGES 219-248
PUBLISHED 23 MARCH 2007**

<http://www.demographic-research.org/Volumes/Vol16/8/>
DOI: 10.4054/DemRes.2007.16.8

Research Material

Evaluation of a village-informant driven demographic surveillance system

Andreas Jahn

Amelia C. Crampin

Judith R. Glynn

Venance Mwinuka

Elenaus Mwaiyeghele

Johnbosco Mwafilaso

Keith Branson

Nuala McGrath

Paul E. M. Fine

Basia Zaba

© 2007 Jahn *et al.*

This open-access work is published under the terms of the Creative Commons Attribution NonCommercial License 2.0 Germany, which permits use, reproduction & distribution in any medium for non-commercial purposes, provided the original author(s) and source are given credit. See <http://creativecommons.org/licenses/by-nc/2.0/de/>

Table of Contents

1	Introduction	220
2	Methods	221
2.1	Setting	221
2.2	Overview of the system	221
2.3	Baseline census	224
2.4	Recruitment and training of village informants	224
2.5	Mapping and formation of clusters and reporting groups	225
2.6	Surveillance of births and deaths	226
2.7	Surveillance of changes of household membership and movements	227
2.8	Data cleaning and consistency checks	228
3	Evaluation	229
4	Results	230
4.1	Continuous demographic surveillance	230
4.2	Evaluation of the sources of event notification and completeness of capturing of events	232
4.3	Completeness of infant death registration	236
4.4	Staffing, logistics and costs	236
5	Discussion	238
6	Acknowledgements	240
	Appendix	245

Evaluation of a village-informant driven demographic surveillance system in Karonga, Northern Malawi

Andreas Jahn ^{1,2}

Amelia C. Crampin ^{1,2}

Judith R. Glynn ²

Venance Mwinuka ¹

Elenaus Mwaiyeghele ¹

Johnbosco Mwafilaso ¹

Keith Branson ²

Nuala McGrath ^{1,2}

Paul E. M. Fine ²

Basia Zaba ²

Abstract

This paper describes and evaluates the first demographic surveillance system (DSS) in Malawi, covering a rural population of 30,000. Unlike others, the Karonga DSS relies on trained village informants using formatted registers for the primary notification of vital events and migrations. Seven project enumerators subsequently collect detailed data on events notified by the village informants, using stringent identification procedures for households and individuals. Internal movements are traced systematically to augment event registration and data quality. Continuous evaluation of data collection is built into the methods. A re-census conducted after 2 years indicated that the routine system had registered 97% of 1,588 births, 99% of 521 deaths and 92% of 13,168 movements.

¹Karonga Prevention Study, PO Box 46, Chilumba, Malawi

²London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, U.K.

1. Introduction

Demographic Surveillance Systems (DSSs) have been established in many developing countries to provide data on fertility and mortality and on determinants of trends in health (9, 18). These systems complement nationally representative surveys such as the Demographic and Health Surveys (DHS) which, along with decennial censuses, are the main source of demographic data in countries lacking universal vital registration systems. DHS has been designed to yield high quality data on fertility, family planning and child health but has been less successful at documenting levels, trends, age patterns and causes of adult mortality. DSS use ‘verbal autopsies’ (VAs), to ascertain the causes of death in broad categories in communities where many deaths occur at home (2, 3, 10, 25). This consists of a structured interview with an informant close to the deceased, with the reported signs interpreted by a physician to infer a diagnosis. If VAs are linked to Sample Vital Registration it may be possible to obtain nationally representative data on causes of death (24).

We initiated the first DSS in Malawi in 2002, in Karonga District in the north of the country. Various models were considered for the new surveillance system, including the widely used Household Registration System pioneered in Navrongo, Ghana (1). Most existing DSSs employ teams of professional enumerators to re-census every household two to four times a year, checking the actual composition against an expected listing of household members. In many DSSs, (e.g. Agincourt, Dar Es Salaam, Hai, Ifakara, Morogoro, Manhiça, Kisesa) the identification of individuals is linked to their relationship to the head of household (14, 19, 26) and this can make it difficult to trace individuals moving between households. The system in Karonga was developed with the specific capacity to identify individuals regardless of their current household and residence, to capture the links between parents and their biological children and to keep household visits by project staff to a minimum to facilitate field logistics and to prevent “interview fatigue” of the study population.

The Karonga DSS was established within the Karonga Prevention Study (KPS), a large epidemiological study of leprosy and tuberculosis which has been running since 1979. This has provided background data from total population surveys conducted in the 1980s, as well as established methods for identification of individuals encountered in health facilities and households and for tracking of individuals enrolled in longitudinal studies (7, 21). The new DSS is based on continuous recording of births and deaths (‘vital events’), and changes of household membership and migrations by trained village informants, with follow-up visits by professional enumerators. The system was evaluated with a re-census after two years. We describe the methods and present the results of the evaluation.

The Karonga DSS, known as the “Continuous Registration System” (CRS) joined the

'International Network for the continuous Demographic Evaluation of Populations and Their Health in developing countries' (INDEPTH) in 2002.

2. Methods

2.1 Setting

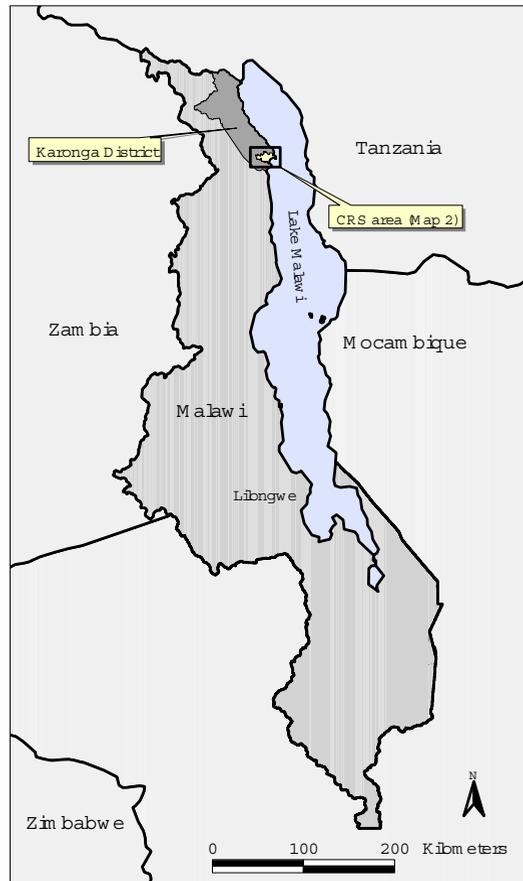
Karonga is a rural district in the Northern Region of Malawi. Approximately 236,000 people live in the District in some 280 villages. The economy is based on subsistence agriculture and fish from Lake Malawi (23). The total population of Karonga District was covered by two whole population surveys from 1979-84 (Leprosy Evaluation Project 1, LEP1) and 1986-89 (LEP2) which were conducted to measure leprosy prevalence and incidence and to recruit individuals for a tuberculosis and leprosy vaccine trial. In the LEP1 survey, 112,000 individuals were enumerated (22) and in LEP2 146,000 (20). Data from these surveys have been used to describe the demography of the district in the 1980s (6). In these surveys and in all subsequent studies by KPS in the district, permanent identifiers have been assigned to all individuals, and people are re-identified on each contact with the project. Details from each contact are added to a continuously maintained database. By the start of the CRS in August 2002, the identifiers of 260,000 people from Karonga District had been accumulated from 600,000 study contacts.

The CRS covers a population of 30,000 in an area of 135km² in the southern part of the District (see Map 1 and Map 2). The area was selected because it was one of two areas used for early studies of HIV prevalence and impact (7, 8, 13), and for logistical reasons (proximity to the project headquarters in Chilumba, and avoiding areas that become inaccessible during the wet season).

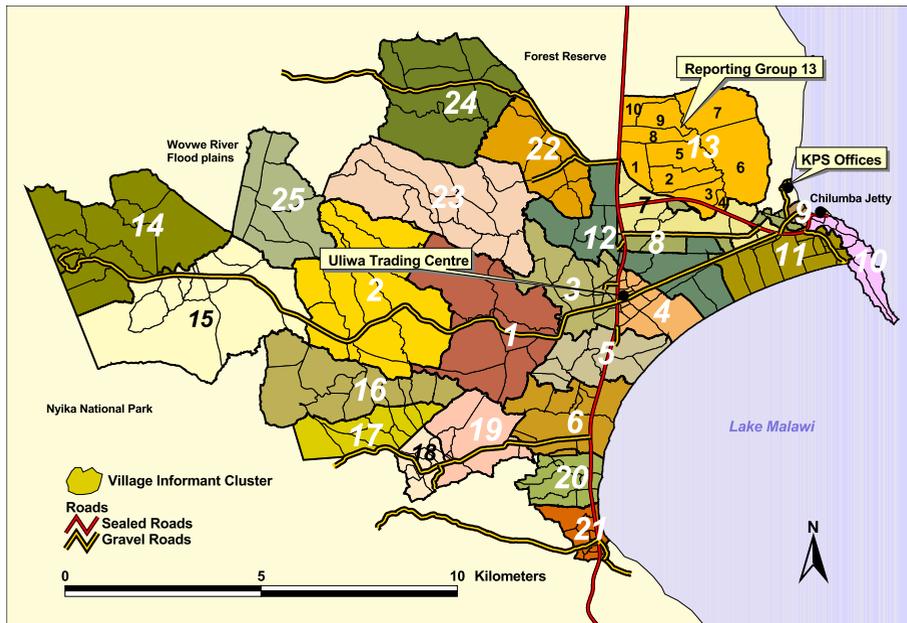
2.2 Overview of the system

An initial house-to-house census was conducted recording personal identifiers and socio-demographic data for all individuals, and economic data and GPS location for every household. The area was mapped and divided into 230 geographically defined *clusters* containing 15-60 households. The clusters were aggregated into 25 *reporting groups* (see Map 2). One *village informant* was sought amongst the residents of each cluster at the time of the census and trained to start the continuous recording of demographic events. Computer-generated household member registers were produced by data clerks in the office and issued to the village informants for the systematic recording of the events. A monthly schedule was set up and convenient venues chosen for meetings between project enumerators and the village informants of each reporting group. At these monthly meetings the project enumerators transcribe the events recorded by the village informants into

Map 1: Malawi, Karonga district and the Karonga CRS area



Map 2: Schematic map of the Karonga CRS area displaying the current 230 village informant clusters grouped into 25 reporting groups (RG; shaded areas). The individual numbering of the clusters is displayed for RG 13.



a computer-generated and manually updated master household register and organise the follow up of households with births or deaths.

Throughout the year, village informants continuously record changes of household membership and migratory movements on a separate log-sheet and once a year, all households in the cluster are visited by the village informants to update the membership list, and report membership changes and migratory movements. Project enumerators visit households with reported changes of membership to record the details. All other household visits by project staff are also used to register opportunistically any previously unreported membership changes ('opportunic registration'), which are recognised by comparing the current status of the household with a master field register. A re-census of the total CRS population was initiated two years after the start, in order to assess the accuracy and completeness of the system.

Ethical approval for these studies was obtained from the National Health Sciences

Research Committee of Malawi and the London School of Hygiene & Tropical Medicine ethical committee.

2.3 Baseline census

The baseline census was conducted by seven project enumerators between August 2002 and July 2004, who operated from camps erected in the respective enumeration areas. Census interviews of each household were carried out at a pre-arranged date. Identifiers and socio-demographic data were collected on each individual and household location, dwelling characteristics and economic data were recorded at the household level. A senior adult member usually acted as the informant for children under the age of 16 or for absent members. Repeat visits were scheduled if a suitable informant was not present or any significant information (specifically relating to the details used for identification) could not be obtained. A detailed description of the identification system for individuals is given in the appendix (section 8.1). As complete enumeration was considered essential for the launching of the CRS, there was no formal limit to the number of repeat visits.

2.4 Recruitment and training of village informants

Each village in the study area has several traditional advisors (“ndunas”) appointed for life by the village headman. Each nduna is responsible for a geographically defined section of the village and supervises all households within that section. The size of the section varies according to the terrain and the population density; typically 20 - 60 households. Ndunas are responsible for broadcasting the news of a death and play an important role in the organisation of funerals. New households and any new members or visitors in existing households have to report to the nduna. This system is still well maintained in rural communities, though it has been discontinued in many urban and road-side communities.

Because of their role as well-respected traditional authorities, ndunas were recruited preferentially as village informants. They were usually well-motivated to take an active role in monitoring demographic events within their community, and most were used to keeping household registers for village development activities. However, approximately 10% of ndunas were illiterate and therefore unable to manage a household register, or they were too frail to visit the households in their area regularly and to attend the monthly reporting sessions. In other cases, the area controlled by a particular nduna covered more than 50 - too many for reliable demographic reporting. In these situations, and whenever a village informant needed to be replaced, the village headman and other ndunas were consulted to appoint a suitable new or additional person as village informant.

Standardised training sessions were run by census enumerators for the village informants of each new *reporting group* (see 2.5) using specially designed posters with illustra-

tions and cartoons, highlighting CRS aims, covering the definition of demographic events and instructions on how to use the household registers. Practical examples were discussed and standardised recording of vital events practised using real events observed in the area. Village informants were advised that they were only expected to visit households that they had not heard anything about for more than one month, as it was anticipated that they would learn about births and deaths by word-of-mouth.

Annual refresher training sessions are held for all reporting groups, repeating the basic training and the rules for the registration of movements are explained in preparation for the annual update session.

The village informants were recruited on a voluntary basis and only a nominal fee is paid to acknowledge the time and effort spent attending each monthly reporting session (see 4.4).

2.5 Mapping and formation of clusters and reporting groups

Within the traditional administrative structures of Karonga District, a 'village' is a group of scattered rural households that declare allegiance to one headman. Villages are not necessarily contiguous geographical areas and the fluidity of the concept of allegiance leads to dynamic groupings of households which can split when the population increases or merge with other villages when political changes occur.

The approximate location of villages was available from household coordinates derived from aerial photographs - these data were reviewed for the initial selection of villages for the CRS. The current village boundaries and the demarcation of the areas controlled by individual ndunas were electronically tracked by GPS using local guides. Although there was not always a recognisable structure on the ground to define existing boundaries, the guides appointed by the village authorities were generally able to identify them with confidence. For the purposes of the CRS, the boundaries established at baseline will be maintained even if the village boundaries move.

Whenever available, village informant clusters were based on the areas controlled by ndunas. Balancing the logistical implications of managing a large number of village informants with small clusters against the expected difficulties for village informants in keeping track of a large number of households in extended clusters, the aim was to allocate 30 households to each village informant. This number was reduced in areas with high migration rates and in sparsely populated areas with scattered households, and increased in 'quieter' rural areas. Whenever new, 'artificial' boundaries had to be defined to delineate more suitably sized clusters, clearly recognisable structures were sought to ensure that the village informants were aware of the area of their cluster.

At the completion of the baseline survey, the median cluster size was 0.43km² (range: 0.01 - 5.19km²) and each cluster contained on average 29 (range: 17 - 60) households. For

logistical purposes, between 7 and 13 adjacent clusters were aggregated into “reporting groups”. These were usually formed to include one or two entire villages. All village informants of a reporting group meet with an enumerator at a central venue for monthly reporting sessions (see Map 2).

The location of every household enumerated in the census was recorded using simple handheld geographical positioning system (GPS) units (Garmin Etrex and Garmin Geko 201). The accuracy was typically within 10m. All main roads, tracks, paths, rivers, health facilities and a selection of public buildings such as churches and schools were located to serve as points of reference for printed maps for field use and for geo-spatial analyses. Coordinates are updated whenever a household moves to a new location or when a new household is formed.

2.6 Surveillance of births and deaths

Village informants record all births and deaths occurring in their cluster in their household register as soon as they become aware of the event. During monthly reporting sessions, the village informants notify the attending enumerator of new births and deaths by showing him the entries in their household registers. The enumerator updates the master household register with the event details and plans the follow up. The enumerator also systematically updates the village informants’ household registers with any details from the master household register recorded during household visits since the previous reporting session.

Households with new births are visited directly after the session by the enumerator, in order to complete the birth registration forms. Deaths are followed up separately by an enumerator trained as a health assistant after a mourning period of at least two weeks. A verbal autopsy (VA) questionnaire is used, based on the standard INDEPTH tool (15), and including details of health care sought and place of death. At every household visit, however triggered, the enumerator compares the members listed in the master register with the current members reported by the household. Any new birth, death or movement thus discovered is registered opportunistically by updating the master register and filling appropriate event forms.

Village informants were not initially asked to report miscarriages and stillbirths, but experience showed they had difficulties distinguishing these events from neonatal deaths which they were supposed to report. Unexpectedly, a large number of miscarriages and stillbirths were readily reported by the village informants and the training protocol was amended to include the routine reporting of these events. The affected households are visited by the VA enumerators to determine and record the nature of the event.

2.7 Surveillance of changes of household membership and movements

The rationale for the surveillance of household membership was to obtain an accurate denominator for the calculation of vital rates, to update information on current residence for all individuals, to facilitate sampling for epidemiological studies, and to analyse the patterns of change of membership status and physical movements of individuals and households. Building on previous studies (5), a household in the CRS is defined by allegiance to a household head and is an independent unit that can change location while maintaining its social identity until dissolution. A household identity is not tied to one head but can be 'inherited' by a new head if s/he explicitly describes the household that way. Only polygamists are permitted to be listed simultaneously (as head) in multiple households. In rare cases, when a member claims allegiance to more than one household, the individual is recorded only in their 'main' household (self-determined by the individual and/or the household).

It is possible for several households to be co-resident in one dwelling if they acknowledge different heads. A household member need not be full time resident; for example school children who reside with other households during school terms are considered to remain members of their home household as long as they regularly return during term breaks.

The recording of household membership as opposed to residence corresponds well with the perception of 'belonging' in this community and simplifies the village informants' decision as to whether events should be reported or not. For instance, it is common for sick individuals to change their residence for short intervals to be nursed by relatives while remaining a member of their original household, but village informants intuitively allocate a death occurring during such an interlude to the usual household rather than the place where the person was nursed.

Definitions of household membership are based on local perceptions; however, specific examples are given during the standardised village informant training session to ensure consistent reporting of changes of household membership. Changes of membership and movements are recorded by the village informants on log-sheets in their household register throughout the year but they are not routinely reported until the update session at the end of an annual reporting cycle. Before this special reporting session, the village informants are asked to visit every household in their cluster to update the member list systematically, including any new households that may have appeared in the cluster. Events recorded on the log-sheets are compared with the current member lists and log sheets are used to record dates of movements and destinations for departures. Annual update sessions are attended by 2-3 enumerators and the group of village informants provides a useful forum to reconcile internal movements and validate unclear events, as village informants are often familiar with households in neighbouring clusters. After the session,

all new households and households with new members or with vital events are visited by the enumerators. As usual, every household visit is used to detect and opportunistically register further unreported events.

In addition to the village informant reporting and opportunistic registration of changes of membership and movements, all individuals moving within the CRS are traced by a visit to the arrival household. These visits are in turn used for opportunistic registration, thus an event in one household can trigger a chain of visits to other households, connected by migrations.

Information about residence status and intervals of absence from the dwelling was collected for every individual during the baseline census and during demographic surveillance for every new household member. A pilot survey had shown that questions specifically addressing where a person had slept the night before the interview were considered intrusive and caused embarrassment. Therefore, questions enquired when an individual had last been at the dwelling using the cut-offs: present during the interview, today, yesterday, 7 days ago and 4 weeks ago.

Many dwellings in Karonga district are not well constructed and have a short life-span. They are commonly abandoned when damaged and new dwellings are erected near the old site. As a consequence, households may move into new dwellings while remaining at approximately the same location. Characteristics of dwellings are recorded at baseline (as a measure of socio-economic status) and are updated if changes occur, but dwelling identifiers and physical labels have not been used in the past or the current studies.

All households present in the surveillance area at the time of the baseline census should have been enumerated. However, given the possibility of missed or newly formed households, village informants were routinely instructed to report any demographic event in their clusters, regardless of whether the event occurred in a household already listed in their files or not.

2.8 Data cleaning and consistency checks

A comprehensive data entry and management system was custom-designed in MS Access 97 to interface with the core modules of the pre-existing FoxPro KPS database. The CRS data supervisor runs a series of batch checks against the database at monthly intervals to verify the consistency of information entered for different individuals and over time for the same individual. Time is split into “demographic episodes” during which an individual is a member of a particular household in a particular cluster. An episode starts at birth, baseline interview, or movement into that household or cluster, and stops at death or movement away from the household or cluster.

Regular checks for inconsistencies include: checks for overlapping demographic episodes, since by definition a household member can be registered in only one household

at a time (with the exception of polygamous heads of multiple households); checks for records of internal movement but with no registration of arrival; checks on correct linkage of birth registration to the mother's details; and checks on households without adults. The fact that no child-headed households have been observed in other studies in Karonga District to date was utilised to check households appearing on the database where the oldest member was less than 18 years of age. This situation occasionally arose if all adults in a household had a departure or death record but the corresponding forms were missed for the remaining minors.

3. Evaluation

A bi-annual re-census of the CRS area is planned, the first re-census was started after 24 months of reporting, with each reporting group undergoing a re-census directly after an annual update session. This allows comparison of events reported at the session with the status of the household members as observed by enumerators during a subsequent household visit. The evaluation presented here is based on the results from the re-census of the first 13 reporting groups (95 clusters), as completed by November 2005. These reporting groups comprise 18,988 (60%) of the 31,646 population estimated to be alive in the CRS area in November 2005.

A special household observation form was used to classify the source of notification for all demographic events registered in the CRS. This records the reason for household observations and flags any events that had not been reported before the household visit and were only discovered through comparison of the list of members in the master register with the current household members.

- routine reporting by the village informants
- household observations triggered by the enumerators tracing internal migrants
- opportunistic registration by the enumerators after comparison of member list in master register or chance discovery of a birth or death in a household
- scheduled visit (e.g. 12-month baby follow-ups - see below, or data consistency queries)
- re-census

For events reported by the village informants, the interval between the first reporting session following the event and the actual reporting date was used to determine whether reporting was timely or late. For events registered through other sources of notification, the dates are analysed to determine whether the village informant would have had the opportunity to report, or whether the enumerators had registered the event early. Events that were registered early have been classified as 'non-missed' by the village informant

and have been excluded when assessing the performance of individual village informants and their clusters.

Babies born in the CRS were visited at 12 months of age to evaluate the completeness of routine infant death registration and to collect data on vaccinations, nutrition and anthropometry. If any of them were found to have died and this had not been previously reported, this event was marked as 'missed' on the household observation form.

The distribution of age at death for children under 2 has been used to estimate the completeness of death registration as the probability of under-reporting is known to be higher for very early infant deaths (17). The distribution of age at death under 2 years observed in the CRS was compared to the national distribution published in the 2000 Malawi DHS (17).

4. Results

4.1 Continuous demographic surveillance

Results from the baseline census are given in the appendix (section 8.2).

At the time of the evaluation in November 2005, 69,000 person years of observation had accumulated in the database. Figure 1 displays the staggered start of the continuous demographic surveillance in all 25 reporting groups and shows the schedule of the monthly and annual reporting sessions and the re-census. As indicated by the blue squares in Figure 1, only the first 13 groups had completed the re-census after 2 years of reporting and could be evaluated.

The 1,588 births and 521 deaths reported in these groups were distributed unequally between the clusters. The median number of vital events (births and deaths) recorded in each informant cluster in 24 months was 20.5, with a range of 6-39 events.

By the end of the re-census of the first 13 reporting groups, 4,300 household visits had been conducted by the enumerators to the 3,791 households resident in the area for the registration of demographic events. Based on 7,000 household years of observation accumulated in these reporting groups, each household had been visited by an enumerator on average 0.6 times per year since the baseline census.

The average time from event to follow-up interview was 4 weeks for births, 5 weeks for deaths and 23 weeks for migrations (Table 1). These average intervals are a result of the village informant reporting cycle (monthly for vital events, annual for migrations) and additional opportunistic registration throughout the year.

Figure 1: Schedule of the staggered start of the reporting groups and their monthly reporting sessions.

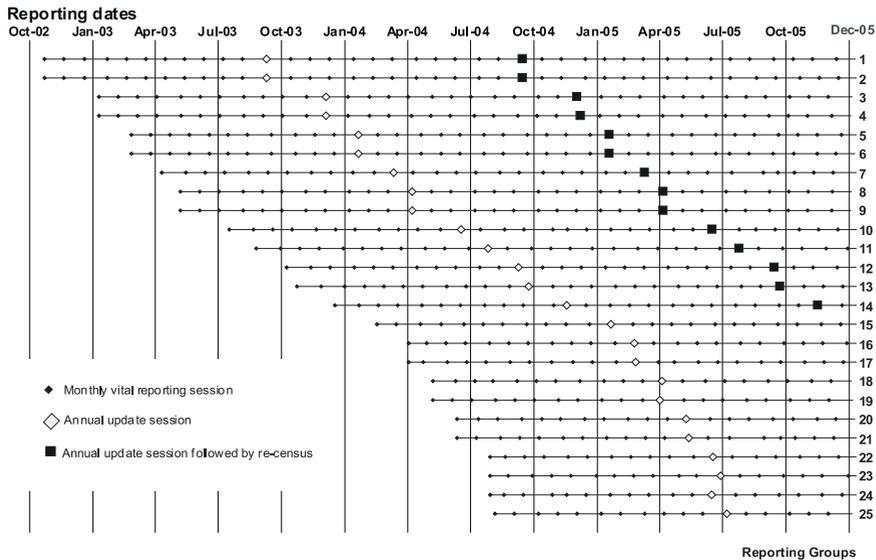


Table 1: Follow-up times from date of the event to the registration visit to the household by the project interviewer

	Median (weeks)	Inter-quartile range
Births	3.8	2.3 - 7.1
Deaths	4.9	3.3 - 7.1
Household change	23.4	11.3 - 37.7

4.2 Evaluation of the sources of event notification and completeness of capturing of events

1,588 births, 521 deaths and 13,168 movements had occurred in the 13 reporting groups by the time of the re-census.

Analysis of the recorded sources of notification showed that 92.2% of the births, 95.0% of the deaths and 61.7% of the movements ultimately recognised were reported by the village informants at the next or later session (combined proportions for category 1 and 2 in Table 2). Excluding the events that were registered ‘early’ by the enumerators (i.e. before the village informant had had the opportunity to report at the next appropriate reporting session) the village informants had ‘missed’ 7.5% of births, 4.4% of under-five deaths, 4.2% of deaths aged five or above, 19.7% of moves in and 20.8% of moves out of households. More than half of these ‘missed’ events were detected by the routine CRS procedures before the re-census (combined categories 7-9 vs. category 10 in Table 2). 17.9% of all movements had been registered opportunistically by the enumerators in the 12-month-intervals between annual reporting sessions for movements.

Overall, the re-census identified 37 births (2.3%), 7 deaths (1.3%) and 1,023 movements (7.8%) that had not (yet) been registered by the routine CRS combination of village informant reporting and opportunistic recording (Table 2).

Table 2: Sources of notification of the demographic events that had occurred before the re-census in the 13 evaluated reporting groups

Sources of notification of events	Births 1,588	Deaths < 5 92	Deaths ≥ 5 429	Moves in 6,554	Moves out 6,614
<i>‘Non-missed’ by village informant*</i>					
1. VI reported at next session	65.6%	76.1%	80.4%	60.5%	57.2%
2. VI reported late	26.6%	17.4%	14.9%	3.3%	2.4%
3. Follow up of internal migration	0.0%	0.0%	0.0%	6.8%	9.0%
4. Opportunistic registration	0.4%	2.2%	0.5%	8.2%	8.4%
5. Query or scheduled follow up	0.0%	0.0%	0.0%	1.4%	2.2%
6. Re-census	0.1%	0.0%	0.0%	0.0%	0.0%
<i>‘Missed’ by village informant*</i>					
7. Follow up of internal migration	1.8%	0.0%	0.9%	2.4%	4.1%
8. Opportunistic registration	3.2%	2.2%	1.9%	6.4%	5.3%
9. Query or scheduled follow up	0.2%	0.0%	0.2%	2.5%	4.3%
10. Re-census	2.3%	2.2%	1.2%	8.4%	7.1%

* An event is classified as ‘non-missed’ by village informant (VI) if it was registered at the household before the VI would have had the opportunity to report at next reporting session. An event is classified as ‘missed’ by VI if it was picked up after a reporting session where the VI failed to report the event.

A considerable proportion of vital events and a small proportion of movements were reported late, i.e. after the first session at which they could have been reported by the village informants (26.0% of all births, 15.6% of all deaths, 2.3% of all movements) and late reporting was more likely if the birth or the movement had occurred within the week before a reporting session (see Table 3); this association was not seen for deaths. It is likely that many of these events classified as ‘missed’ by the village informant would have been reported at a later session. However, since enumerators did not record reports that were made after an event had already been registered opportunistically, it is impossible to quantify this.

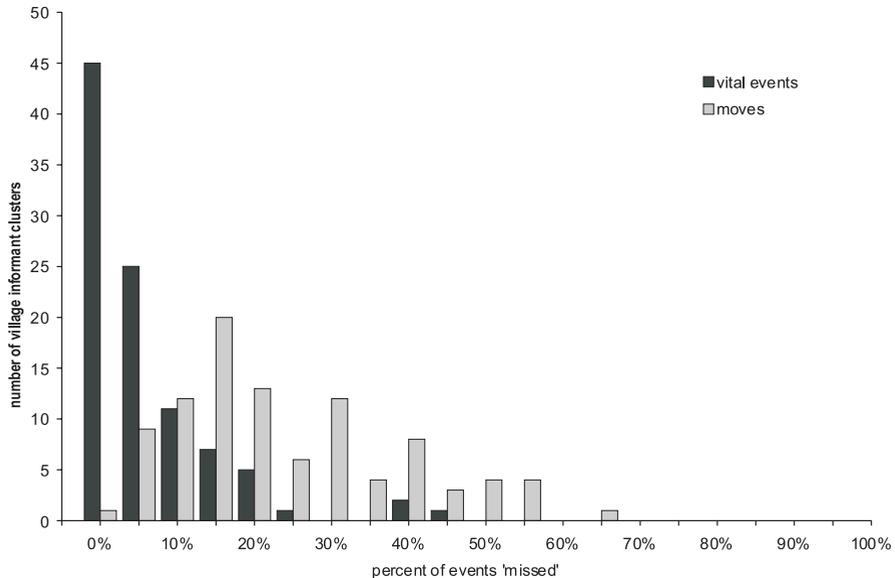
Table 3: Timing of event in relation to the next reporting session. The table shows the proportions of events that were reported by village informants at the next opportunity (excluding events registered ‘early’)

	n	reported at next session	p-value for diff. of prop.
Births			
All	1,581	65.5%	
Event 1-7 days before next session	364	51.4%	
Event ≥ 8 days before next session	1,217	69.7%	<0.001
Deaths <5 years			
All	90	77.8%	
Event 1-7 days before next session	20	90.0%	
Event ≥ 8 days before next session	70	74.3%	0.9
Deaths ≥ 5 years			
All	427	80.6%	
Event 1-7 days before next session	87	78.2%	
Event ≥ 8 days before next session	340	81.2%	0.2
Moves into households			
All	5,470	64.4%	
Event 1-7 days before next session	168	50.6%	
Event ≥ 8 days before next session	5,302	64.8%	<0.001
Moves out of households			
All	5,315	66.7%	
Event 1-7 days before next session	140	55.7%	
Event ≥ 8 days before next session	5,175	67.0%	0.002

The distribution of the 96 evaluated village informants by the proportions of ‘missed’ events is displayed in Figure 2. In one of these clusters the village informant had to be replaced approximately half-way through the evaluation period due to poor performance

and both village informants are included in the analysis; the standard errors were adjusted for the non-independence of the cluster characteristics.

Figure 2: Distribution of the village informant-clusters by the proportions of ‘missed’ vital events and moves.



The proportion of movements which were ‘missed’ was generally higher (median 21.3%) than that of vital events (median 5.3%) in all clusters. In order to analyse the factors associated with village informants ‘missing’ events in their clusters, a critical threshold of 10% was defined for ‘missed’ vital events and 20% for ‘missed’ movements. The associations are displayed in Table 4. There was some indication that female village informants were less likely to ‘miss’ both vital events and movements, but this was not statistically significant. Older village informants were more likely to miss movements (borderline significant) but not births and deaths. Village informants’ educational attainment was weakly associated with missing fewer migrations ($p=0.2$), but it did not appear to influence the proportions of ‘missed’ vital events. While ndunas performed significantly worse in reporting movements than non-ndunas, this association was not seen for the reporting of births and deaths. The poorer performance of ndunas was explained by their older age and the fact that they had on average more extended clusters with more

households and a lower population density, factors that were independent predictors for more ‘missed’ movements (see below). After taking account of these characteristics, there was no significant difference between ndunas and other village informants with regards to more than 20% missed movements (OR: 1.54 [0.60-3.98]).

Table 4: Association of village informant (VI) and cluster characteristics with ‘missed’ events

Village informant characteristics		n	proportion of VI-clusters above threshold of			
			≥10% ‘missed’ vital ev.		≥20% ‘missed’ moves	
VI gender	male	82	30%		60%	
	female	15	13%	p=0.2	40%	p=0.2
VI age	<40	19	26%		42%	
	40-49	31	19%		45%	
	50 +	47	34%	p=0.7 †	70%	p=0.06 †
VI education*	primary 1-4	19	42%		68%	
	primary 5-8	52	17%		58%	
	sec. or higher	25	40%	p=0.9 †	48%	p=0.2 †
VI is nduna	no	45	20%		44%	
	yes	52	35%	p=0.1	67%	p=0.02
Cluster characteristics						
number of households in cluster	<25	18	22%		39%	
	25-34	32	44%		59%	
	35-44	26	12%		46%	
	45-54	14	29%		79%	
	55 +	7	29%	p=0.6 †	86%	p=0.02 †
cluster size (km ²)	<0.25	32	28%		41%	
	0.25 –	22	32%		50%	
	0.5 –	24	33%		79%	
	≥1	19	16%	p=0.5 †	63%	p=0.02 †
pop. density (persons/km ²)	<250	24	21%		67%	
	250-499	31	29%		61%	
	500-749	13	38%		62%	
	750 +	29	28%	p=0.6 †	41%	p=0.06 †

* educational level was not recorded for one VI

† test for linear trend across ordered groups

While there was no clear trend in the relationship between the number of households per cluster and the proportion of 'missed' vital events, village informants with more households were significantly less likely to report more than 80% of the movements in their cluster. Interestingly, there appeared to be a critical threshold of 45 households per cluster below which the completeness of reporting of movements did not vary significantly. A similar association was seen for the cluster size (area in km²), with an increased proportion of movements being 'missed' in clusters of 0.5km² or larger. Conversely, a higher population density was associated with a lower risk of 'missed' movements above the threshold. Cluster area or population density showed no clear association with the proportions of vital events 'missed'. There was no indication that the total number of vital events or movements per cluster influenced the proportion of events 'missed' (results not shown).

4.3 Completeness of infant death registration

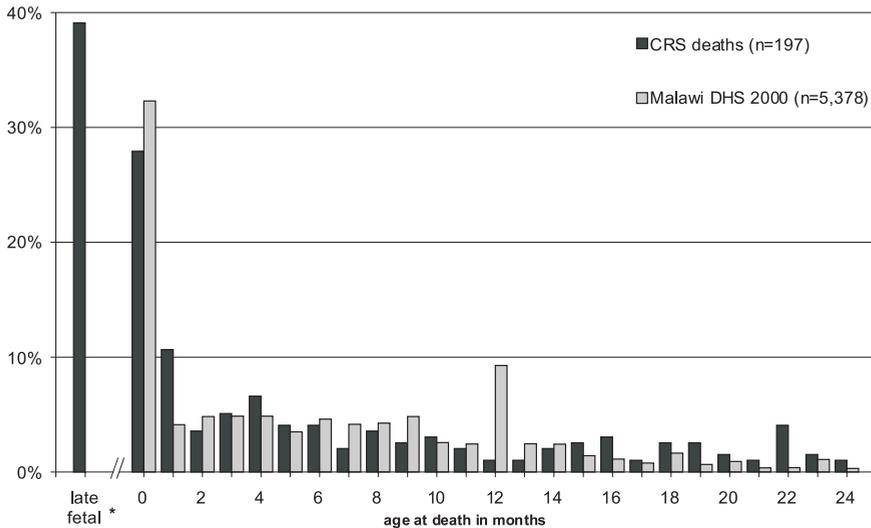
662 babies, who were born in one of the 13 reporting groups and who were believed to be still alive and resident in the area, were visited at 12 months. This scheduled follow-up identified only one infant whose death had not been reported or registered otherwise.

The distributions of the 197 deaths under 2 years of age registered in the CRS as of November 2005 and the national estimates from the 2000 Malawi DHS are displayed by month in Figure 3; 37.2% of infant deaths captured in the CRS had occurred within the first month of life compared to 42% estimated in the 2000 Malawi DHS. The proportion of neonatal deaths that occurred in the first week of life was 67.6% in the CRS and 67% in the 2000 DHS (distribution by weeks not shown in Figure 3) (17). Figure 3 also displays the 77 late foetal deaths (miscarriages after ≥ 28 weeks of gestation and stillbirths) registered in the CRS. The ratio of late foetal deaths to deaths within the first 7 days of life (early neonatal deaths) was 83:51 (1.63) in the CRS and 148:116 (1.28) in a large community follow-up study of pregnant women in the south of Malawi in 1996 (16).

4.4 Staffing, logistics and costs

The Karonga CRS was set up as part of a multi-disciplinary research programme and therefore shared the existing project infrastructure and core administrative and support staff. Demographic data collection currently relies on a pool of 230 trained village informants for the continuous registration of demographic events in the community and a team of 7 full-time professional enumerators who organise the monthly reporting sessions and conduct all household visits to fill detailed event registration forms and update field registers. The monthly village-informant-fees total approximately \$7,000 annually. All project enumerators are equipped with motorbikes and cover a total of 5,500km per month

Figure 3: Distribution of age at death in months comparing the events registered in the CRS up to November 2005 with the estimates from the 2000 MDHS derived from retrospective birth histories.



*) Estimates for late fetal deaths are not available from the 2000 MDHS; the proportion of late fetal deaths registered in the CRS is calculated using the total of the under 2 deaths as the denominator.

to attend reporting sessions and visit households in the surveillance area of 135km². Five trained data-clerks and one supervisor are responsible for the processing, entry and data cleaning of the approximately 15,000 incoming event questionnaires per year and for the production and checking of the field registers. The CRS was designed, implemented and supervised by the authors, all of whom were involved to a varying degree in other scientific activities. Their salaries have not been included in the costing, but the equivalent of a full time demographer/epidemiologist is needed to run the study.

To estimate the total annual running costs, a proportion of the core project costs has been added to the CRS budget. With the current staffing levels (excluding the scientist) and logistical setup, an average of approximately \$2.4 is spent on each individual under surveillance per year.

5. Discussion

Approximately 60 Demographic Surveillance Systems have been established in developing countries worldwide over the past 50 years and 33 have joined the INDEPTH network. Though the designs of these projects differ, most have been able to provide high quality demographic and health data. The Karonga CRS was designed to generate detailed longitudinal demographic data and to provide an accurate sampling frame for epidemiological studies. Building on the existing traditional structures in this rural African community, the system was designed to maximise the participation of respected community members in the data collection process while minimising the annoyance to the population which may arise through repeated household enumeration. The Karonga CRS requires an average of 0.6 visits per household per year by project staff and is less intrusive in the community than those requiring yearly or more frequent visits.

We found that our village informants could be trained to keep a standard pre-formatted household register and to capture vital events reliably although most of them had only a very basic educational level. The analysis of village informant and cluster characteristics suggested that females and individuals under the age of 50 were more successful in the registration of births and deaths and that clusters should be designed to comprise no more than 45 households and to cover an area of not more than 0.5km² for reliable reporting of movements by village informants. The performance of the village informants was only partly predicted by their socio-demographic characteristics, and the individual motivation and reliability of reporting varied. The systematic evaluation of the routine reporting and registration system in the re-census led us to review the recruitment criteria for village informants and the formation of household clusters. Twenty-one of the original 209 clusters were subsequently split and a number of older informants (mainly *ndunas*) were replaced.

The re-census two years after the start showed that only a small proportion of births and deaths had escaped the routine system of village informant reporting and opportunistic registration. Considering the observed proportions of events reported with delay it is likely that fewer than 3% of births and 2% of deaths would not have been picked up eventually without the re-census. Although these small proportions would not greatly influence the estimation of fertility and mortality rates, the accumulation of missed events without periodic re-censuses would gradually weaken the population database used for epidemiological sampling so it appears that a two-yearly re-census cycle is appropriate. The large pool of village informants, like any team of permanently employed enumerators, requires regular supervision and quality checking to detect possible fatigue or protocol drift and continuous monitoring and regular evaluation is essential for a long-running system like the CRS.

The comparison of the distribution of age at death in children under 2 years recorded in the CRS with the estimates from the 2000 DHS and the data from the community

follow-up-study (16) suggests that there was no selective omission of late foetal, perinatal and infant deaths in the CRS. Given that the events most likely to be missed are the very early infant deaths, this may indicate that the death registration across other age-bands was nearly complete, Early neonatal deaths may be confused with late foetal deaths (16) and the slightly different ratio from that found in the other study from Malawi (16) may be due to misclassification in retrospective birth history interviews.

Permanent identifiers for the primary subjects under demographic surveillance (individuals and households) are a prerequisite for the generation of consistent and high quality DSS data. The Karonga CRS utilises a system for the identification of individuals and households developed by the project 20 years before the start of the DSS - a system that is independent from the allocation of individuals and households to residential units. As shown from previous studies, the reliability of the identification process enables us to track individuals and households over long periods of time even if they return from outside the district (4, 6-8). It also allows linkage of information from related studies in the same population (including outside the CRS area), and does not require the individual to be present each time identification is required (unlike an ID card or biometric system).

The particular strengths of the Karonga CRS, compared to other studies, lie in its detailed data on migrations, including the linkage of individuals over time as they migrate within, in and out of the area. Births and deaths are followed up within a very short interval as they do not have to wait for the next census round. The linkage of children to both parents and the linkage of individuals to their spouses, even if they do not share a household, is a unique resource for genetic studies and for studies of orphanhood and childcare arrangements. Community informants may have more insight into and knowledge of dynamics such as changes of households than outside enumerators. The disadvantages of the infrequent census cycles are that socio-economic data are less frequently updated, and collecting additional data on the population requires the organisation of extra surveys.

DSS data collection methods are costly and the long-term continuous funding of the sites remains a challenge. Due to the comparatively small number of employed field staff that are required for the event-triggered registration system, the annual budget required to run the Karonga CRS is modest.

The methods developed for the Karonga CRS may well be applicable to other areas, particularly rural areas where traditional community structures remain strong.

6. Acknowledgements

This work is supported by grants from The Wellcome Trust, UK and the British Leprosy Relief Association (LEPRA). JRG is supported in part by the UK Department of Health. We would like to thank Sian Floyd for their helpful comments and the National Health Sciences Research Committee of Malawi for their continued support.

References

- (1) Binka F.N., Ngom P., Phillips J.F., Adazu K.F., MacLeod B.B. (1999): "Assessing population dynamics in a rural African society: The Navrongo demographic surveillance system." *J Biosoc Sci.*, 31:375-91.
- (2) Chandramohan D., Maude G.H., Rodrigues L.C., Hayes R.J. (1994): "Verbal autopsies for adult deaths: issues in their development and validation." *Int J Epidemiol.*, 23(2):213-22.
- (3) Chandramohan D., Maude G.H., Rodrigues L.C., Hayes R.J. (1998): "Verbal autopsies for adult deaths: their development and validation in a multicentre study." *Trop Med Int Health.*, 3(6):436-46.
- (4) Chirwa T. (2001): "Effect of household dynamics on risk of disease associated with household contact" [PhD]. London: London School of Hygiene & Tropical Medicine.
- (5) Chirwa T., Floyd S., Ponnighaus J., Malema S., Kileta S., Zaba B., et al. (2004): "Household dynamics in northern Malawi during the 1980s." *Southern African Journal of Demography.*, 9(2): 1-23.
- (6) Chirwa T., Floyd S., Zaba B., Fine P.E. (2005): "Demographic analysis of the population of Karonga District, northern Malawi, in the early years of the HIV epidemic (1980-90)." (*submitted*).
- (7) Crampin A.C., Floyd S., Glynn J.R., Sibande F., Mulawa D., Nyondo A., et al. (2002): "Long-term follow-up of HIV-positive and HIV-negative individuals in rural Malawi." *AIDS.*, 16(11):1545-50.
- (8) Crampin A.C., Floyd S., Glynn J.R., Madise N., Nyondo A., Khondowe M.M., et al. (2003): "The long-term impact of HIV and orphanhood on the mortality and physical well-being of children in rural Malawi." *AIDS.*, 17(3):389-97.
- (9) D'Souza S. (1981): "A population laboratory for studying disease processes and mortality—the Demographic Surveillance System, Matlab Comilla." *Rural Demogr.*, Bangladesh, 8(1):29-51.
- (10) Datta N., Mand M., Kumar V. (1988): "Validation of causes of infant death in the community by verbal autopsy." *Indian J Pediatr.*, 55(4):599-604.
- (11) Fitness J., Floyd S., Warndorff D.K., Sichali L., Malema S., Crampin A.C., et al. (2004): "Large-scale candidate gene study of tuberculosis susceptibility in the Karonga district of northern Malawi." *Am J Trop Med Hyg.*, 71(3):341-9.

- (12) Fitness J., Floyd S., Warndorff D.K., Sichali L., Mwaungulu L., Crampin A.C., et al. (2004): "Large-scale candidate gene study of leprosy susceptibility in the Karonga district of northern Malawi." *Am J Trop Med Hyg.*, 71(3):330-40.
- (13) Glynn J.R., Ponnighaus J., Crampin A.C., Sibande F., Sichali L., Nkhosa P., et al. (2001): "The development of the HIV epidemic in Karonga District, Malawi." *AIDS.*, 15(15):2025-9.
- (14) INDEPTH Network (2002): *Part III. INDEPTH DSS site profiles. Population and health in developing countries Population, health and survival at INDEPTH sites.*, Ottawa, Canada: International Development Research Centre; p. 127-211.
- (15) INDEPTH Network (2003): *INDEPTH Standardized Verbal Autopsy questionnaire*(Revised August 2003), Available from: http://www.indepth-network.org/core_documents/indepthtools.htm [cited 2005 03/10/2005]
- (16) McDermott J., Steketee R.W., Wirima J.J. (1996): "Perinatal mortality in rural Malawi." *Bull World Health Organ.*, 72(2):165-71.
- (17) Ndawala J.S. (2001): "Infant and child mortality." *In: National Statistics Office [Malawi], ORC Macro, editors. Malawi Demographic and Health Survey 2000.*, Zomba, Malawi and Calverton, Maryland, USA: ORC Macro; 2001. p. 97-104.
- (18) Ngom P., Binka F.N., Phillips J.F., Pence B., Macleod B. (2001): "Demographic surveillance and health equity in sub-Saharan Africa." *Health Policy Plan.*, 16(4):337-44.
- (19) NIMR Mwanza Centre (2005): *Kisesa Cohort Study TANESA Program (MAGU DSS). MAGU DSS Procedures.*, Available from: <http://www.magudss.net/projects/maduprocedures.htm> [cited 2005 03/10/2005]
- (20) Ponnighaus J.M., Fine P.E., Bliss L., Gruer P.J., Kapira-Mwamondwe B., Msosa E., et al. (1993): "The Karonga Prevention Trial: a leprosy and tuberculosis vaccine trial in northern Malawi. I. Methods of the vaccination phase." *Lepr Rev.*, 64(4):338-56.
- (21) Ponnighaus J.M., Fine P.E., Sterne J.A., Bliss L., Wilson R.J., Malema S.S., et al. (1994): "Incidence rates of leprosy in Karonga District, northern Malawi: patterns by age, sex, BCG status and classification." *Int J Lepr Other Mycobact Dis.*, 62(1):10-23.
- (22) Ponnighaus J.M., Fine P.E., Bliss L., Sliney I.J., Bradley D.J., Rees R.J. (1987): "The Leptra Evaluation Project (LEP), an epidemiological study of leprosy in Northern Malawi." *I. Methods. Lepr Rev.*, 58(4):359-75.

- (23) Republic of Malawi PD (2002): *Karonga District Socio Economic Profile. Chapter I. General Features.*, Lilongwe; 2002.
- (24) Setel P.W., Sankoh O., Rao C., Velkoff V.A., Mathers C., Gonghuan Y., et al. (2005): "Sample registration of vital events with verbal autopsy: a renewed commitment to measuring and monitoring vital statistics." *Bull World Health Organ.*, 83(8):611-7.
- (25) Snow R.W., Armstrong J.R., Forster D., Winstanley M.T., Marsh V.M., Newton C.R., et al. (1992): "Childhood deaths in Africa: uses and limitations of verbal autopsies." *Lancet.*, 340(8815):351-5.
- (26) Tollman S.M., Herbst K., Garenne M., Gear J.S., Kahn K. (1999): "The Agincourt demographic and health study - site description, baseline findings and implications." *S Afr Med J.*, 89(8):858-64.
- (27) Wallace C., Fitness J., Hennig B., Sichali L., Mwaungulu L., Ponnighaus J.M., et al. (2004): "Linkage analysis of susceptibility to leprosy type using an IBD regression method." *Genes Immun.*, 5(3):221-5.

Appendix

Identification of individuals

In this population few people have personal identification documents and personal details (including names) may vary at successive interviews. Names are typically spelt according to pronunciation and therefore inconsistently, and it is common for a person to ‘discard’ previously used names and to assume new names over the course of his or her life, often in the context of specific demographic events (e.g. adulthood, marriage, divorce).

Although educational levels have improved since the 1980s, an appreciable proportion of the population, particularly women and older people, are unable to provide a precise year of birth. The project originally developed a local events calendar to estimate the period when a person was born when the informant was unable to report a precise year (22). The events calendar was later replaced by birth decade estimates.

Due to these factors, personal identifiers recorded on successive interviews do not always match exactly the identifiers of the same individual stored in the database. The correct re-identification thus relies on simultaneous comparison of multiple characteristics to determine a match.

The characteristics routinely recorded for identification are:

- Names (family names, given names, nicknames, clan names and additional names that may have been used in the past)
- Sex
- Date of birth
- Mother’s and father’s details (names, survival status, name of current household head, village)
- Village of birth
- Names of up to 3 full- or half-siblings alive in other households in the district
- Name of current spouse(s)
- Name of current household head and current village of residence
- Name of previous household head and previous village of residence
- Year and location of any previous contacts with the project, as recalled by the respondent

This detailed identification system was developed during the LEP surveys in the 1980s. Although a formal evaluation has not been conducted due to the unavailability of an independent system that could be used as a ‘gold-standard’, experience has shown that individuals are reliably re-identified even when the interval between the contacts is more than a decade. For example, over the course of 20 years, 430,000 study contacts with previously known individuals required re-identification from the database, and

records indicate that in only 0.7% of repeat contacts had known subjects not been correctly recognised and assigned with a new identifier.

An advantage of this identification system compared to those based on bio-features such as photographs or fingerprints is that it is usually possible to ascertain the identity of a household member from a proxy informant in the household without the need for each individual to be present during the interview.

Since 1981, the identifiers of both the biological mother and the biological father have been collected systematically for every individual recruited, allowing the reconstruction of complete family trees. These data have been used for genetic studies of tuberculosis and leprosy. (11, 12, 27)

Baseline census

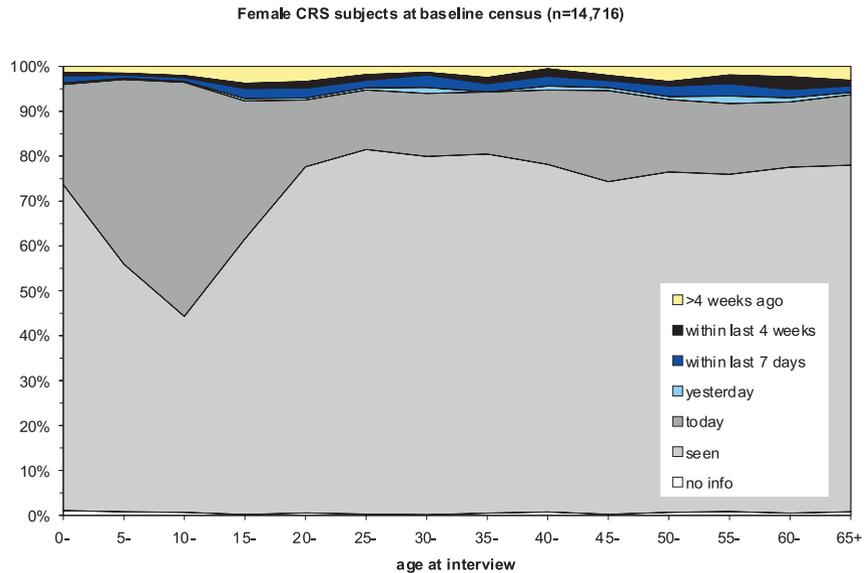
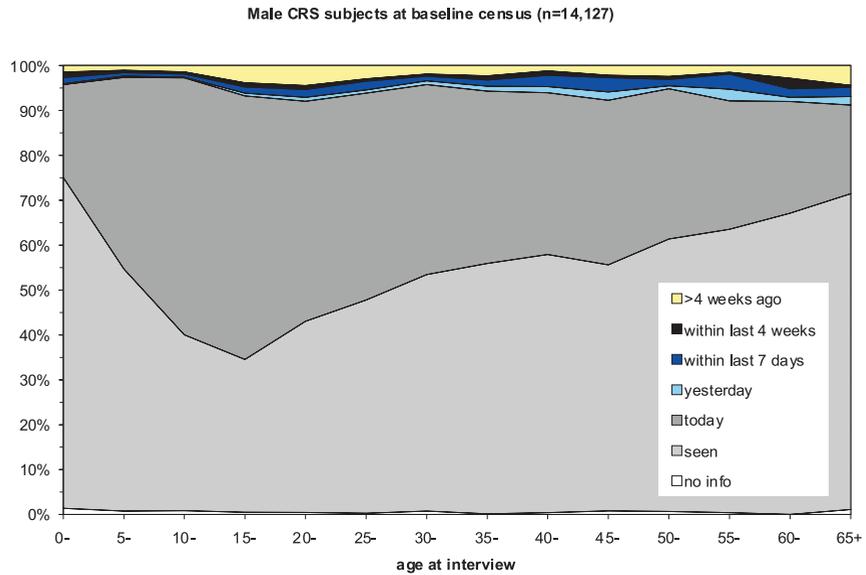
The study has been well received by the community. In the course of the baseline census and the following continuous surveillance, the field team encountered fewer than 25 households that initially refused to be interviewed. Only a small community of Old Apostolic Believers (approximately 15 households) completely refused to participate, on the grounds of their faith, and have had to be excluded permanently.

The baseline census enumerated 28,843 individuals and 5,820 households. 988 (17%) households were headed by polygamists and although each polygamist was only counted as one individual he was recorded as the head in each of his households. In total, 5,274 individuals (18.3%) were classified as heads (of 1 or up to 5 households), 23,526 (81.6%) as regular members, 38 (0.1%) as renters living with a resident household and 5 (0.02%) as itinerants (i.e. homeless individuals without permanent affiliation to any household). A further 1,218 individuals were classified as non-members: 95.3% of these were reported to be visitors, 4.2% as household employees and for 0.5% the position was missing. Visitor- and employee-status was only assigned if the person was considered to have maintained their membership in another household. Due to their non-membership status, these individuals were excluded from surveillance in the household where they were originally enumerated.

The median reported duration of stay for visitors was 3 months (50% of the visit intervals were between 1 and 13 months) and 95.4% of visitors had been at the household on the day of the interview.

Figure 4 displays the proportion of household members actually seen at their dwelling as recorded during the baseline census by age and sex. Although an appointment had been made with each household during the booking visit, encouraging adults to be present for the census interview, only 60.8% of household members (53.8% of males and 67.7% of females) were actually seen at the census visit. A further 33.4% were recorded to have been at the household on the day of the interview but had left before the enumerator

Figure 4: Pattern of presence and absence from the dwelling of CRS subjects at the baseline census.



arrived. Thus, a total of 94.2% of household members had been present on the day of the census visit. Only 2.3% of 14,623 males and 2.1% of 14,934 females who were included as household members had not been at the dwelling during 4 weeks prior to the census interview and might therefore be considered not currently resident. The highest proportions of members not present within the last 4 weeks were found in the 15-24 year olds (males: 4.1%; females: 3.6%) and in members aged 50 or above (males: 3.2%; females: 2.8%).

Of the 4,832 non-polygamist heads of household (males and females combined), 64.8% were seen at their household, whereas 30.6% were not seen but had been at the dwelling on the day of the interview, 2.6% had not been there on the day of the interview but had been within the last 4 weeks and 2.0% had been away for more than 4 weeks. The distribution was very similar for the 451 polygamist heads. Thus household heads were less likely to be long-term absent than were household members.

To date approximately 60 households were discovered that were reported to have been present in the CRS area at the time of the baseline census but that had escaped enumeration. These missed households were commonly brought to the attention of the enumerators by the report of a demographic event through a village informant and this notification triggered a household visit. For analyses, any birth or death that occurred before the first household interview that initiated the continuous demographic observation was considered left-censored and excluded.

