

An assessment of household deaths collected during Census 2011: A discussion document

**Prepared to be released with the 10% sample of Census 2011 data
April 2014**

**Embargoed until:
15 April 2014**

Enquiries:

Christine Khoza
+2771 670 2672

Forthcoming issue:

Expected release date

15 April 2014 10:00 AM

Table of contents

1. Introduction	2
2. Preliminary evaluation of samples of death records	4
3. The error resolution exercise	6
4. Results of the manual capturing exercise	9
4.1 Type of captured values	10
5. Records deleted after the application of the modified final edit specification rules	12
6. A comparison of age ratios of deaths	13
7. A comparison of total number of deaths between household and vital register deaths	15
8. Conclusion	16
9. Acknowledgements	16
References	17
Appendix	18

Tables

Table 1.1: Distribution of household deaths by province: Published Tables, Census 2011	4
Table 1.2: Distribution of household deaths by province: Final data	4
Table 2: Number of manually captured death records by province	10
Table 3: Number of death records changed	11
Table 4: A comparison of household deaths between published tables and final data (unweighted)	11
Table 5: A comparison of age ratios between household and vital register deaths	14

Figures

Figure 1.1: An example of an error resolution interface for the month of death	7
Figure 1.2: An example of a blank death record with images	8
Figure 1.3: An example of a blank death record with a line drawn through person number eight	9
Figure 2: A comparison of total number of deaths over time	15

1. Introduction

The collection of household deaths data has been found to be problematic in most developing countries, in spite of international efforts to use censuses as a vehicle to solicit this information (UN, 2008). Reasons for this include under-reporting (the most affected groups are children and the elderly), reference-period errors and the unwillingness of the respondents to talk about recent deaths (Timæus, 1991). Another cited shortcoming is the exclusion of deaths due to the dissolution of households after the death of a breadwinner. In the case of South Africa, the inclusion of questions on household deaths in the Census2011 questionnaire was the second attempt following Census 2001. According to Dorrington & Moultrie (2004), analyses of the 2001 household deaths revealed that the month and the year of death as well as the age of the deceased were subjected to some unnecessary editing (3,1%, 6% and 8% respectively).

For the Census2011, questions on household deaths were asked in SECTION I: MORTALITY IN THE LAST 12 MONTHS in Questionnaire A. The descriptions of questions pertaining to household deaths are available in the Census 2011 METADATA at <http://www.statssa.gov.za>. These questions were administered to populations living in households. Persons living in institutions were excluded. However, persons who were found living in households within an institutional area (e.g. caretakers) were included as they were enumerated with Questionnaire A. Fieldwork monitoring reports by the Monitoring and Evaluation division within Statistics South Africa (Stats SA) revealed that Section I did not do well from the Pilot to the main census. Some reasons for under-reporting were:

- Fieldworkers were afraid to ask the questions or skipped these deliberately, the reason being that some respondents felt those were too personal and emotional questions.
- Some questions were left blank.

All in all, a combination of respondent fatigue, sensitivity of the mortality questions, poor training strategy and lack of supervision during data collection appear to be the underlying factors for under-reporting when one looks at the responses on household deaths.

The release of Census 2011 results one year after the data collection exercise was well received by stakeholders during the dissemination workshops held across the country. That notwithstanding, the ongoing coding of industry, occupation and migration as well as assessments of both the fertility and household deaths data led to the delay of the 10% sample release. With regard to published mortality tables, about 20% of the total number of deaths had unspecified values for both age and sex. This feature became a cause for concern, as

presented in Table 1.1. These death records were ultimately flagged for further investigation at the time of the release. Such death records include those for which both the month and the year of death are not stated, an invalid month (not 1–12) and not stated year of death, year of death not stated, invalid or out of the reference period (10/2010 to 10/2011). One reason for including such records at the time of the release was the suspicion of erroneous reading by the scanning system.

Apart from death records that were flagged for further investigation, a few deaths were reported as having occurred in November and December 2011. A number of respondents called the call centre to request enumeration after 31 October 2011. As a result, enumeration continued up to early December 2011. About 5 489 (weighted) death records had November and December 2011 reported for the month and the year of death, constituting less than one per cent of the total number of deaths. Analysis of the comparison of deaths by month of death between the vital register and those enumerated during Census 2001, as well as those enumerated during the 2007 Community Survey, reveals that some of the enumerators tended to put the month of enumeration as the month of death. It was considered therefore to accept such cases as in-scope by evenly distributing them throughout the reference period months from October 2010 to October 2011 in the edited data.

Table 1.1 and Table 1.2 present a comparison of Census 2011 household deaths between Published Tables (Census 2011 Statistical Release: P0301.4) and those tabulated from the final data (expected to be released with the 10% sample). Lower values of deaths are shown in Table 1.2 relative to those shown in Table 1.1 for all provinces, although there is a variation regarding the magnitude. Provinces that reflect higher percentages of reduced number of deaths include Gauteng and Western Cape (32% and 25% respectively), whilst Eastern Cape shows the lowest percentage (14%). The most affected cases are those whose age and sex were unspecified at the time of the release. The detailed process that produced these changes is explained below.

Table 1.1: Distribution of household deaths by province: Published Tables, Census 2011

	Total number of deaths	Unspecified age	Unspecified sex	Unspecified age and sex	Unspecified age	Unspecified sex	Unspecified age and sex
Province	Number				Per cent		
Western Cape	45 453	849	160	10494	1,9	0,4	23,1
Eastern Cape	92 185	2 354	355	12 359	2,6	0,4	13,4
Northern Cape	14 369	386	53	2 130	2,7	0,4	14,8
Free State	44 318	881	153	8 315	2,0	0,3	18,8
KwaZulu-Natal	136 636	7 948	843	21 292	5,8	0,6	15,6
North West	45 903	1 202	153	7 832	2,6	0,3	17,1
Gauteng	118 066	9 983	650	27 427	8,5	0,6	23,2
Mpumalanga	51 828	1 046	161	13 656	2,0	0,3	26,3
Limpopo	55 786	885	172	14 715	1,6	0,3	26,4
South Africa	604 544	25 534	2 700	118 220	4,2	0,4	19,6

Source: Statistics South Africa

Table 1.2: Distribution of household deaths by province: Final data

	Total number of deaths	Unspecified age	Unspecified sex	Unspecified age and sex	Unspecified age	Unspecified sex	Unspecified age and sex
Province	Number				Per cent		
Western Cape	34 026	163	160	3	0,5	0,5	0,0
Eastern Cape	79 330	1 986	271	72	2,5	0,3	0,1
Northern Cape	12 033	338	43	15	2,8	0,4	0,1
Free State	35 535	720	92	18	2,1	0,3	0,1
KwaZulu-Natal	109 768	4 018	521	174	3,7	0,5	0,1
North West	37 705	1 035	94	32	2,7	0,2	0,1
Gauteng	80 839	2 095	288	87	6,2	0,4	0,1
Mpumalanga	37 944	903	117	24	2,6	0,3	0,1
Limpopo	40 887	748	112	33	1,8	0,3	0,1
South Africa	468 067	12 006	1 698	458	2,6	0,4	0,1

Source: Statistics South Africa

2. Preliminary evaluation of samples of death records

The purpose of this undertaking was to do further investigation into household deaths flagged as such at the time of the release. With no preconceived notions of what constituted blanks for both age and sex, a need arose to first view a sample of scanned questionnaire images of all aforementioned death records. Of the total of 518 002 unweighted death records (not adjusted for the undercount) available for screening, about 115 221 matched the status of being flagged for further investigation. About 653 of the 115 221 death records were randomly selected for screening. The results of the screening showed that, during processing, about 21% of the 653 selected records were erroneously read by the scanning system as valid records, whereas the

response to whether a death had occurred in the household during the reference period was “No”.

In the event that a line/s or some figure-like images having been created on the questionnaire image emanating from dirt, these translated into figures for at least one mortality variable in the raw data (see an example of a blank record with figure-like images and one with a line drawn through a specific case in Figure 1.2 and Figure 1.3). About 4% of the screened records were out of the reference period (October 2010 to October 2011). However, among the aforementioned records, some were erroneously declared as out of scope just by erroneously reading the month or year of death differently than what is reported on the questionnaire. For example, when the month and year of death are reported as April 2011 on the scanned questionnaire image, but read as April 2010 in the raw data, this pushes it out of the reference period, whilst it was actually reported within said period. Others were declared as out of scope due to the enumerator failing to utilise the year of birth of the deceased provided by the respondent to derive the age of the deceased, in the event that it was not known by the respondent. The year of birth would also have been written on some parts of the questionnaire, probably as a means to enable the calculation of age at some stage in the process of enumerating. In almost all cases, whether the age of the deceased was calculated correctly or incorrectly by the enumerator, the year of birth would be put in as the year of death even after calculating the age of the deceased.

Some of such records were ultimately included back into the raw data after the incorporation of manually captured cases. This was done by summing up the year of death (which in this case would be the year of birth of the deceased) and the reported age in the event that the year of death was out of scope and the reported month valid, which would then amount to either 2010 or 2011 as the year of death. Although such cases were minimal, a few were actually lost as a result of the enumerator’s erroneous calculation of the age of the deceased, say by one year higher or lower than the actual age implied. This led to a slightly exaggerated number of out-of-scope death records.

In contrast to the 653 death records flagged for further investigation, the results of about 900 randomly selected death records from the remaining 402 781 death records that were not flagged for further investigation showed consistency between the scanned questionnaire images and the raw data. Although some may argue that both screening samples were rather

too small, some insights pertaining to data problems being linked to death records flagged for further investigation at the time of the release assisted in deciding the way forward.

A decision was made therefore, to manually capture all 115 221 death records. Figure 1.1 gives an example of the error resolution interface for the month of death. The exercise focused on the four key household death variables, namely month of death, year of death, sex, and age of the deceased. Given this, each of the 115 221 cases would have four error resolution interfaces developed for the four variables respectively, amounting to 460 884. This does not mean that the other variables were free of errors. The modified final edit specification rules were expected to deal with the remaining data problems upon the incorporation of manually captured cases into the bigger raw mortality data. For example, the minimum processability rule would resolve cases where the pregnancy-related responses were erroneously captured by the scanner reading system. For example, if a value was erroneously provided for whether the deceased was pregnant at the time of death, emanating from dirt or other scanner related problems, that record would not meet the minimum processability rule.

3. The error resolution exercise

Error resolution involves the use of the scanned questionnaire to verify missing or out-of-range values in the raw data. It is usually done after scanning, but before data editing. In the case of Census2011, such an exercise was done following the initial data capturing through scanning. However, it would appear there was a challenge regarding the mortality variables, where for most of the blank variables, these could not be verified since there were no values on the scanned images either. Some of the cases that were blank for the month and the year of death were changed to being unspecified (99 and 9999 respectively). The error resolution term may actually be misleading in this context, as a decision was made to manually capture all four variables as they appear on the screened image. This meant that even when all the four variables reflected unacceptable values (missing, out of range or unclear) on the screened image, these were captured as they appeared.

About 11 employees were seconded from the Processing division within Statistics South Africa to augment the three members of the mortality monograph. All these employees were trained on the error resolution interface by the system developers. The first manual capturing phase ran for about three weeks. Upon the completion of the first phase, it emerged that about 84 197 of the manually captured 460 884 values in the data were not consistent with the scanned

questionnaire images. Since the principle used was to capture what is on the questionnaire image, regardless of being out of scope, invalid or not stated, it was decided that there is a need to resolve those inconsistent cases. The second manual capturing phase focused on the aforementioned 84 197 cases. About eight data processors were tasked to do some quality assurance on those remaining inconsistent cases.

Figure 1.1: An example of an error resolution interface for the month of death

SECTION I: MORTALITY IN THE LAST 12 MONTHS

M-00 DEATH OCCURRED
Has any member of this household passed away in the last 12 months (between 10 October 2010 and 9 October 2011)?

☒ 1 Yes
☐ 2 No
☐ 3 Do not know

Mark the appropriate circle with an X.

M-00a NUMBER OF DEATHS
How many members of the household passed away in the last 12 months (between 10 October 2010 and 9 October 2011)?

0 3 3

ASK ONLY ABOUT DECEASED WOMEN THAT WERE AGED 12 - 50 AT THE TIME OF DEATH

M-01 NAME OF DECEASED
What was the first name of (the deceased)?

SINE
THEM
BA

M-02 MONTH AND YEAR OF DEATH
What was the MONTH and the YEAR of (the deceased's) death?

07
2010

M-03 SEX OF THE DECEASED
Was (the deceased's) male or female?

☒ 1 Male
☐ 2 Female

M-04 AGE OF THE DECEASED
What was (the deceased's) age in completed years at the time of death?

0 3 3

M-05 NATURAL OR UNNATURAL DEATH
Was the death due to a natural or an unnatural cause?

☒ 1 Natural
☐ 2 Unnatural
☐ 3 Do not know

M-06 PREGNANT AT TIME OF DEATH
Did (the deceased) die while pregnant?

☒ 1 Yes
☐ 2 No
☐ 3 Do not know

M-07 DEATH DURING BIRTH
Did (the deceased) die while giving birth?

☒ 1 Yes
☐ 2 No
☐ 3 Do not know

M-08 POSTNATAL DEATH
Did (the deceased) die within 6 weeks after delivery?

☒ 1 Yes
☐ 2 No
☐ 3 Do not know

THANK YOU FOR YOUR CO-OPERATION

Shortcuts
Next Field - Ctrl + N
Back Field - Ctrl + B
Rotate - Ctrl + R
Resolve Type - Ctrl + Number

QN Barcode: A0695695830

Page Number: A13

Person Number: 1

Field Name: M02_DEATH_MONTH

First Capture:

Second Capture:

Choose from the above values or type below

Type Value:

Confirm Value:

Resolve Type: 3 - RESOLVED WITH BLANKS

Back

1 - RESOLVED
2 - BY PASSED (NOT CLEAR)
3 - RESOLVED WITH BLANKS
4 - INCORRECT/INVALID PAGE
5 - EXCEED MAX DIGITS
6 - BY PASSED (MULTIPLE VALUES)
7 - BY PASSED (VALUE N/A)

Note: The scanned image shows that the reading system read the month of death as blank since it pointed above the value (07) for the month of death.

Figure 1.2: An example of a blank death record with images

Data Processing Management System (Version = 2.0.0.116) - [User = Jabulane Monaiwa (jabumo) - Server = Production]

Home » Questionnaire Image Viewer

0.46x
1:1
Full

SECTION 1: MORTALITY IN THE LAST 12 MONTHS

M-00 DEATH OCCURRED
Has any member of this household passed away in the last 12 months between 10 October 2010 and 9 October 2011?

☐ 1 Yes
☐ 2 No
☐ 3 Do not know

Mark the appropriate circle with an X.

M-01 NAME OF DECEASED
What was the first name of the deceased?

Use CAPITAL LETTERS only

M-02 MONTH AND YEAR OF DEATH
What was the MONTH and the YEAR of the deceased's death?

Mark the appropriate circle with an X.

M-03 SEX OF THE DECEASED
Was the deceased male or female?

1 = Male
2 = Female

M-04 AGE OF THE DECEASED
What was the deceased's age in completed years at the time of death?

Write the age in the boxes. If age is less than 1 year, write DEC.

M-05 NATURAL OR UNNATURAL DEATH
Was the death due to a natural or an unnatural cause?

1 = Natural (e.g. accident, illness)
2 = Unnatural (e.g. murder, suicide)
3 = Do not know

Mark the appropriate circle with an X.

M-06 PREGNANT AT TIME OF DEATH
Did the deceased die while pregnant?

1 = Yes
2 = No
3 = Do not know

Mark the appropriate circle with an X.

M-07 DEATH DURING BIRTH
Did the deceased die while giving birth?

1 = Yes
2 = No
3 = Do not know

Mark the appropriate circle with an X.

M-08 POSTNATAL DEATH
Did the deceased die within 9 weeks after delivery?

1 = Yes
2 = No
3 = Do not know

Mark the appropriate circle with an X.

ASK ONLY ABOUT DECEASED WOMEN THAT WERE AGED 12 - 50 AT THE TIME OF DEATH

If 1 to M-08 or M-08, Questionnaire completed

If more than 8 deaths in the household, use a second questionnaire. Write the barcode of the 1st questionnaire below:

THANK YOU FOR YOUR CO-OPERATION

Barcode A0760468841

A13

Note: Figure-like images that may have emanated from dirt, as the mortality section is on the back page of the questionnaire, which may translate into figures in the raw data.

Table 2: Number of manually captured death records by province

Province	Total death records manually captured	Records changed	Proportion changed	Number not changed	Proportion not changed
Western Cape	343	342	100	1	0
Eastern Cape	13 281	5 710	43	7 571	57
Northern Cape	2 184	1 292	59	892	41
Free State	8 151	5 353	66	2 798	34
KwaZulu-Natal	25 529	13 212	52	12 317	48
North West	7 382	4 252	58	3 130	42
Gauteng	32 156	21 197	66	10 959	34
Mpumalanga	12 349	8 675	70	3 674	30
Limpopo	13 846	9 984	72	3 862	28
South Africa	115 221	70 017	61	45 204	39

Source: Statistics South Africa

4.1 Type of captured values

Table 3 presents two parts, namely A and B. Part A shows the two sets of manually captured death records. These sets include those that were made blank for all mortality variables and those for which at least one variable was populated with some value. Broadly, about two-thirds of the total 70 017 death records that changed were resolved with a blank for all four variables, as presented in Table 3 (see images of blank death records in Figure 1.2 and Figure 1.3). This finding suggests that a substantial number of cases initially captured as a death record without age or sex information were in fact not death records at all. Table 4 shows a higher number of death records deleted due to simply being blank for all mortality variables (54 334) compared to about 47 800 (a subset of the changed records) presented in Table 3. The extra records shown in Table 4 involve some of the 45 204 death records that were accurately read by the scanning system as such, as shown in Table 2. Provinces that reflect above 70% for such records include Limpopo, Mpumalanga, Gauteng and Free State. Provinces that reflect below 50% for such records include Eastern Cape and KwaZulu-Natal, as presented in Table 3.

Part B shows the disaggregation of the remaining third of the total death records (22 217) that were changed because the variable changed. The sex of the deceased was the most resolved of the four variables whilst the age of the deceased was the least resolved of the four, as shown in Table 3. This stage marks the end of the error resolution phase. The next phase would be data editing for all mortality data.

Table 3: Number of death records changed

A						B							
Province	Total records changed	Made blank for all four variables		At least one variable populated		Month of death		Year of death		Sex of deceased		Age of deceased	
		N	Proportion	N	Proportion	N	Proportion	N	Proportion	N	Proportion	N	Proportion
Western Cape	342	200	58	142	42	108	32	120	35	135	39	45	13
Eastern Cape	5 710	2 474	43	3 236	57	2 545	45	2 652	46	3 116	55	2 233	39
Northern Cape	1 292	769	60	523	40	441	34	439	34	504	39	341	26
Free State	5 353	3 885	73	1 468	27	1 186	22	1 241	23	1 416	26	1 038	19
KwaZulu-Natal	13 212	6 039	46	7 173	54	5 236	40	5 498	42	6 795	51	4 479	34
North West	4 252	2 744	65	1 508	35	1 218	29	1 256	30	1 448	34	1 013	24
Gauteng	21 197	16 109	76	5 088	24	4 064	19	4 280	20	4 860	23	3 672	17
Mpumalanga	8 675	7 118	82	1 557	18	1 237	14	1 336	15	1 499	17	1 089	13
Limpopo	9 984	8 462	85	1 522	15	1 863	19	1 874	19	2 155	22	1 654	17
South Africa	70 017	47 800	68	22 217	32	17 898	26	18 696	27	21 928	31	15 564	22

Notes: Row proportions were derived to provide insight into the magnitude of data errors at provincial level. Total numbers of death records resolved by variable do not necessarily add up to the total provincial number since more than one variable would have been resolved with a value for one record.

Source: Statistics South Africa

Table 4: A comparison of household deaths between published tables and final data (unweighted)

Province	Number of deaths at the time of the release (unweighted)	Month and year of death not stated	Out-of-scope and invalid cases	Not meeting the minimum processability rule	Total records deleted	Final data (to be released with the 10% sample) (unweighted)
Western Cape	37 115	4 629	3 836	856	9 321	27 794
Eastern Cape	81 678	2 742	7 045	1 553	11 340	70 338
Northern Cape	12 401	766	934	276	1 976	10 425
Free State	38 842	4 113	2 741	814	7 668	31 174
KwaZulu-Natal	115 976	6 780	12 267	3 377	22 424	93 552
North West	38 064	2 753	2 980	853	6 568	31 478
Gauteng	99 545	16 657	11 595	2 821	31 073	68 472
Mpumalanga	44 345	7 159	3 512	838	11 509	32 836
Limpopo	50 036	8 735	3 432	905	13 072	36 964
South Africa	518 002	54 334	48 342	12 293	114 969	403 033

Source: Statistics South Africa

The incorporation of the manually captured cases into the raw data entailed super-imposing the new information on the existing one for those barcodes. Thereafter, the modified final edit specification rules were applied to the overall data. The minimum processability rule accepted cases that had either age or sex and other two mortality variables stated. It also stated that in the event that both age and sex were not stated, but the record had the month and the year of death stated within the reference period, it should be accepted as a valid record.

The reader is reminded that of the 518 002 unweighted deaths (not adjusted for the undercount) at the time of the release, only 115 221 records were availed for the manual capturing process. Given this, about 402 781 records were considered to be acceptable. The comparability of the 114 969 records deleted from the release data after the aforementioned intervention regarding the 115 221 manually captured records may prompt an argument as to why the aforementioned error resolution exercise was done rather than just deleting the suspect 115 221 death records to save time. This outcome, although almost the same as the original number of unspecified cases for both age and sex at the time of the release, was driven by multiple changes as highlighted in Table 4.

The reader should at this stage be reminded that the number of deleted records presented in Table 4 refers to the overall raw mortality data after the application of the final modified edit specification rules. The status of the year of death was used as the yardstick for accepting records by virtue of providing the reference period. Cases that may have met the minimum processability owing to either age or sex and two other variables reported would be accepted if two of the reported variables included the month and the year of death, failure to which these would be deleted.

5. Records deleted after the application of the modified final edit specification rules

More than a half of the deleted records (60 631) were deleted due to year of death being out of scope or failure to meet the minimum processability rule, as shown in Table 4. This finding suggests that there were more errors introduced during the data collection phase than those introduced during the processing phase. Records deleted due to being out of the reference period (October 2010 to October 2011) constitute about 42% of the total records deleted, as shown in Table 4. A subset of about 289 death records of records deleted due to being out of scope was deleted due to year of death being invalid (zero to 3-digit values). Out-of-scope values for the year of death include zero, one to 3-digit values with a range of 0–8838.

Records that were deleted due to not meeting the minimum processability rule constitute about 11%. Such records had the year of death either not stated or unspecified. A subset of about 5 625 of such cases was deleted due to missing year of death, while the month of death may either be valid (1–12), invalid or unspecified. Another subset of about 4 508 of such cases was deleted due to the year of death missing and at most two other mortality variables reported. The

remaining subset of about 2 156 involved cases that had unspecified values for both the month and the year of death.

Nearly half of the total death records (54 334) deleted are linked to the month and the year of death not stated, as presented in Table 4. Such records are as a result of data scanning problems. All in all, Gauteng appears to reflect a higher number of records deleted for missing month and year of death, relative to other provinces. However, for records deleted due to failure to meet the minimum processability rule, as well as being out of scope, KwaZulu-Natal is leading.

6. A comparison of age ratios of deaths

Table 5 presents a comparison of numbers of household deaths between the Census 2011 Published Tables (available in the Census 2011 Statistical Report at <http://www.statssa.gov.za>) and the final data (to be released with the 10% sample). The value of the error resolution exercise is evident in the proportion of records deleted due to being unspecified for age and sex (92%) as shown in Table 5. Vital register deaths were tabulated from the beginning of October 2010 to the end of September 2011 (2011 registered deaths are provisional since they are yet to be published) to enable comparison with the Census 2011 reference period.

Table 5: A comparison of age ratios between household and vital register deaths

Age	Census 2011 Published Tables: available on the Census 2011 Statistical Release (P0301.4)	Census 2011 Final Mortality Data (to be released with the 10% sample)	*Vital Register Deaths from the beginning of October 2010 to the end of September 2011	Census 2011 Published Tables: available on the Census 2011 Statistical Release (P0301.4) and Final Mortality Data (to be released with the 10% sample)	Published Tables (available on the Census 2011 Statistical Release (P0301.4) and *Vital Register Deaths from the beginning of October 2010 to the end of September 2011	Census 2011 Final Mortality Data (to be released with the 10% sample) and *Vital Register Deaths from the beginning of October 2010 to the end of September 2011
	Frequency			Ratios		
0	41 635	40 978	28 772	1,0	1,5	1,4
1-4	13 838	13 605	10 331	1,0	1,4	1,3
5-9	6 153	5 812	4 321	1,1	1,4	1,4
10-14	4 294	4 103	3 966	1,0	1,1	1,0
15-19	8 186	8 094	7 740	1,0	1,1	1,1
20-24	19 173	18 920	17 863	1,0	1,1	1,1
25-29	31 546	31 260	32 009	1,0	1,0	1,0
30-34	37 023	36 696	38 469	1,0	1,0	1,0
35-39	37 698	37 496	40 572	1,0	0,9	0,9
40-44	32 415	32 140	37 089	1,0	0,9	0,9
45-49	31 136	30 998	36 278	1,0	0,9	0,9
50-54	29 502	29 250	35 669	1,0	0,8	0,8
55-59	28 383	28 076	34 525	1,0	0,8	0,8
60-64	28 373	28 117	35 384	1,0	0,8	0,8
65-69	23 703	23 492	30 928	1,0	0,8	0,8
70-74	24 686	24 493	33 099	1,0	0,7	0,7
75-79	20 342	20 166	27 680	1,0	0,7	0,7
80-84	18 203	18 076	26 668	1,0	0,7	0,7
85-89	11 462	11 385	16 679	1,0	0,7	0,7
90+	13 038	12 904	15 563	1,0	0,8	0,8
Unspecified	143 755	12 006	2 284	12,0	67,1	5,6
Total	604 544	468 067	515 889	1,3	1,2	0,9

*Vital register deaths for 2011 are published in the 'Mortality and causes of death 2011: findings from the death notification forms' (P0309.2)

Source: Statistics South Africa

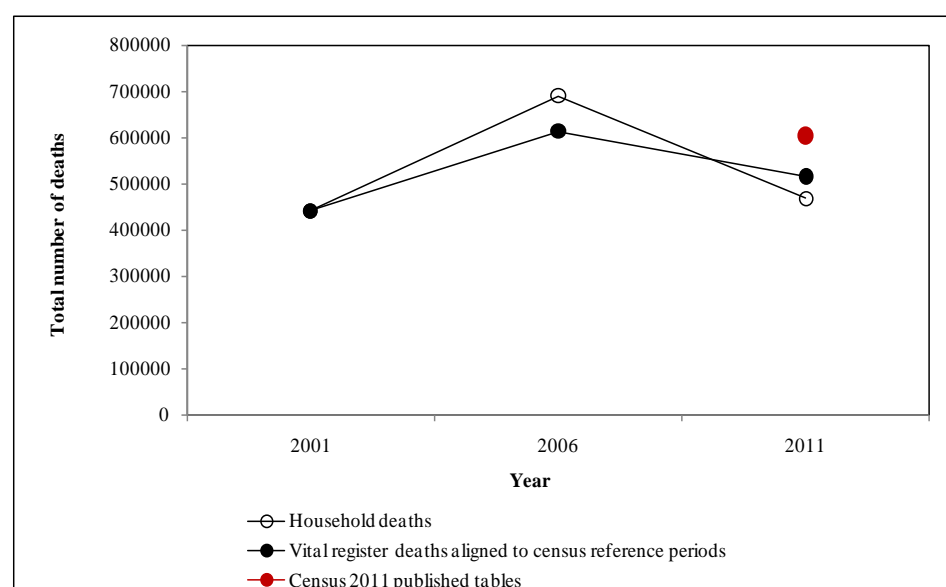
Age ratios of household deaths enumerated during Census 2011 reflect compatibility at all ages. This finding suggests that the age-specific death rates derived are comparable. Age ratios of both the published tables versus vital register deaths as well as deaths tabulated from the final mortality data versus vital register deaths show a similar pattern. Household deaths show higher values for children aged 0–9 relative to the vital register. In contrast, vital register deaths show higher values for adults aged 35 years and above relative to household deaths. This age pattern comparison is expected, given that rural populations may not see the necessity of registering dead children, while the registration of adult deaths links to the administration of estates in the case of South Africa.

7. A comparison of total number of deaths between household and vital register deaths

Figure 2 presents a comparison of the total number of deaths between the vital registration system and enumerated deaths over time. Vital register deaths are aligned with the census/survey reference period. For example, registered deaths for 2001 are tabulated from the beginning of October 2000 to the end of September 2001 for the sake of comparing deaths pertaining to the same period. Also, those for 2006 are tabulated from the beginning of February 2006 to the end of January 2007 and so on.

On the one hand, the total number of household deaths enumerated during the 2007 Community Survey appears to be higher than that provided by the vital register at the same period. On the other hand, Census 2011 household deaths appear to be slightly under-reported relative to Census 2001. Nonetheless, sampling and non-sampling errors cannot be ruled out regarding the 2007 Community Survey. Overall, both the vital register and enumerated data reveal the same upward and thereafter downward trend. That notwithstanding, Census 2011 household deaths implies a steeper decline in the mortality rate relative to vital register deaths at the same period. In essence, the steeper downward trend reflected between the 2007 Community Survey and Census 2011 signifies some over- as well as under-enumeration of household deaths.

Figure 2: A comparison of total number of deaths over time



Note: Vital register deaths for 2011 are published in the 'Mortality and causes of death 2011: findings from the death notification forms' (P0309.2)

8. Conclusion

Although the assessment of household deaths is actually the first of its kind since the onset of mortality data collection through censuses and surveys in the case of South Africa, some of the data problems observed provide some useful insights into what may have been happening during previous attempts. Also, death records that lack year of death should have been discarded earlier on, since keeping them in the data proved not to be useful in the long run. Overall, this exercise revealed that there were actually more data collection errors compared to processing errors in the Census 2011 mortality data. There is a need therefore, for an improved data collection strategy, mainly around data collection particularly pertaining to training strategies in preparation for future censuses.

9. Acknowledgements

The contribution made by the four mortality monograph team members is acknowledged. The names of the aforementioned team members are: Mercy Shoko, Lesego Lefakane, Kadisha Ramaru and Sheila Dlamini. Also, the support provided by staff members within both the Data Processing and the Population Statistics divisions at Statistics South Africa is well acknowledged.

Enquiries:

Name: Christine Khoza

Email: christinek@statssa.gov.za

References

Dorrington, R., Moultrie T.A. & Timæus I.M. 2004. Estimation of mortality using the South African Census 2001 data, Cape Town.

Statistics South Africa. 2012. Census 2011 Statistical Release P0301.4. Available at <<http://www.statssa.gov.za>>.

Statistics South Africa. 2004. Household deaths tables for Census 2001. Pretoria: Statistics South Africa.

Statistics South Africa. Monitoring and Evaluation: Census 2011. Pretoria: Statistics South Africa. (unpublished).

Statistics South Africa. 2013. Mortality and causes of death in South Africa: findings from death notification. Statistical Release P0309.3. Available at <<http://www.statssa.gov.za>>.

Statistics South Africa. 2007. The 2007 Community Survey. Pretoria: Statistics South Africa. Available at <<http://www.statssa.gov.za>>.

Timæus, I. 1991. Measurement of adult mortality in less developed countries: a comparative review. *Population Index* 57(4), pp. 552-568.

United Nations, 2008. Principles and Recommendations for Population and Housing Censuses: Revision 2. New York: United Nations.

Appendix

The attached Appendix replicates the tables that were included in the Census 2011 release by using the revised final figures.

Table 1: Death occurred

	Western Cape	Eastern Cape	Northern Cape	Free State	KwaZulu-Natal	North West	Gauteng	Mpumalanga	Limpopo	South Africa
Yes	32 166	72 199	11 240	33 211	99 641	34 971	74 779	35 483	38 961	432 650
No	1 598 605	1 612 447	289 613	788 523	2 435 052	1 024 772	3 824 209	1 038 153	1 377 070	13 988 444
Do not know	3 154	2 697	547	1 551	4 644	2 256	9 839	1 830	2 054	28 571
Total	1 633 925	1 687 344	301 400	823 285	2 539 336	1 061 998	3 908 826	1 075 466	1 418 085	14 449 665

Note: Total number of households includes housing units and converted hostels only, differences may be due to rounding off.

Table 2: Number of deaths

	Western Cape	Eastern Cape	Northern Cape	Free State	KwaZulu-Natal	North West	Gauteng	Mpumalanga	Limpopo	South Africa
1	31 744	69 946	10 685	31 589	93 939	33 282	73 681	33 746	37 505	416 118
2	1 969	8 178	1 154	3 356	13 075	3 782	6 016	3 637	3 046	44 213
3	246	862	117	418	1 704	452	840	340	228	5 206
4	56	238	59	113	813	160	207	168	82	1 897
5	0	61	5	22	96	17	37	18	17	271
6	0	45	6	19	118	13	41	27	8	277
7	11	0	0	10	0	0	16	0	0	37
8	0	0	8	8	23	0	0	8	0	47
Total	34 025	79 330	12 033	35 535	109 767	37 705	80 838	37 944	40 887	468 065

Note: Differences may be due to rounding off.

Table 3: Distribution of household deaths by age and sex nationally

Age	Male	Female	Unspecified	Total
0	21 885	18 812	279	40 976
1-4	7 248	6 305	52	13 605
5-9	3 074	2 720	18	5 813
10-14	2 194	1 893	16	4 103
15-19	4 279	3 792	23	8 093
20-24	9 251	9 634	35	18 920
25-29	14 939	16 253	68	31 260
30-34	18 671	17 950	74	36 695
35-39	20 324	17 090	82	37 496
40-44	17 747	14 332	61	32 139
45-49	17 530	13 399	70	30 999
50-54	16 627	12 552	71	29 250
55-59	16 326	11 708	43	28 077
60-64	15 965	12 073	79	28 117
65-69	13 024	10 409	59	23 492
70-74	12 488	11 945	60	24 493
75-79	9 380	10 735	52	20 168
80-84	7 568	10 467	42	18 077
85+	8 381	15 855	55	24 292
Unspecified	6 121	5 424	457	12 003
Total	243 023	223 348	1 696	468 067

Table 4: Distribution of household deaths by age, sex and province

Age	Western Cape				Eastern Cape				Northern Cape			
	Male	Female	Unspecified	Total	Male	Female	Unspecified	Total	Male	Female	Unspecified	Total
0	1 232	1 012	44	2 289	3 043	2 610	39	5 692	477	432	10	919
1-4	319	215	8	542	1 134	924	9	2 067	193	127	1	320
5-9	184	132	7	323	518	454	2	975	99	41	0	140
10-14	115	86	0	201	353	309	3	666	51	42	0	93
15-19	357	183	3	543	944	745	6	1 694	116	86	0	202
20-24	867	501	0	1 369	2 028	1 814	9	3 850	212	197	0	408
25-29	983	711	3	1 697	2 835	3 052	15	5 902	315	329	1	644
30-34	967	664	10	1 641	3 314	3 191	6	6 512	414	371	1	787
35-39	1 155	749	7	1 911	3 415	2 854	16	6 286	426	408	0	834
40-44	991	845	10	1 845	2 799	2 228	8	5 036	474	378	4	855
45-49	1 179	874	11	2 065	2 874	2 181	11	5 065	492	385	2	879
50-54	1 365	1 016	3	2 384	2 829	2 112	15	4 956	506	393	3	902
55-59	1 526	1 098	5	2 628	2 632	1 925	3	4 560	458	337	0	795
60-64	1 639	1 199	13	2 852	2 620	2 068	12	4 700	452	377	2	831
65-69	1 469	1 121	5	2 595	2 064	1 813	9	3 886	396	328	2	725
70-74	1 425	1 219	11	2 655	2 314	2 164	14	4 492	365	363	0	728
75-79	1 245	1 143	5	2 392	1 665	1 946	11	3 622	301	305	0	606
80-84	948	961	1	1 910	1 344	1 890	7	3 242	171	243	0	413
85+	788	1 221	12	2 021	1 476	2 662	3	4 141	211	400	2	614
Unspecified	79	81	3	163	980	933	72	1 986	170	153	15	338
Total	18 832	15 033	160	34 025	41 182	37 877	271	79 330	6 298	5 693	43	12 033

Table 4 continued: Distribution of household deaths by age, sex and province

Age	Free State				KwaZulu-Natal				North West			
	Male	Female	Unspecified	Total	Male	Female	Unspecified	Total	Male	Female	Unspecified	Total
0	1 608	1 445	9	3 061	5 985	5 122	73	11 179	1 914	1 715	11	3 639
1-4	500	450	3	954	1 926	1 720	17	3 663	601	531	2	1 135
5-9	214	188	0	402	815	734	3	1 552	241	226	1	468
10-14	131	128	0	259	661	545	8	1 214	174	115	0	290
15-19	284	253	1	539	1 064	1 055	6	2 125	293	261	1	555
20-24	561	706	1	1 268	2 496	2 779	11	5 286	547	757	2	1 306
25-29	1 086	1 210	4	2 300	4 421	4 352	24	8 797	897	1 172	1	2 070
30-34	1 438	1 418	2	2 858	5 194	4 434	22	9 650	1 452	1 354	10	2 816
35-39	1 603	1 447	4	3 054	5 152	3 993	27	9 172	1 733	1 566	8	3 307
40-44	1 534	1 372	6	2 911	4 039	3 074	12	7 124	1 573	1 220	3	2 796
45-49	1 453	1 262	4	2 720	3 788	2 744	25	6 557	1 474	1 159	3	2 636
50-54	1 403	1 060	9	2 472	3 322	2 467	18	5 807	1 369	1 099	4	2 472
55-59	1 204	991	4	2 200	3 279	2 276	17	5 572	1 359	994	1	2 354
60-64	1 246	1 003	7	2 256	3 353	2 562	17	5 932	1 297	863	6	2 165
65-69	947	750	5	1 703	2 528	2 188	13	4 729	1 095	774	0	1 869
70-74	847	825	4	1 676	2 302	2 662	17	4 981	1 035	807	1	1 843
75-79	610	784	1	1 395	1 584	2 314	16	3 915	827	777	3	1 607
80-84	428	672	3	1 104	1 374	2 474	11	3 859	650	711	1	1 362
85+	535	1 143	4	1 682	1 472	3 154	9	4 635	711	1 267	2	1 980
Unspecified	351	352	18	720	2 049	1 795	174	4 018	565	437	32	1 035
Total	17 985	17 458	92	35 535	56 803	52 445	521	109 768	19 805	17 806	94	37 705

Table 4 continued: Distribution of household deaths by age, sex and province

	Gauteng				Mpumalanga				Limpopo			
Age	Male	Female	Unspecified	Total	Male	Female	Unspecified	Total	Male	Female	Unspecified	Total
0	3 457	2 930	44	6 431	2 086	1 735	32	3 853	2 084	1 811	19	3 913
1-4	1 145	1 011	5	2 160	696	670	4	1 370	734	656	2	1 392
5-9	483	462	2	948	281	243	0	524	239	240	2	481
10-14	290	286	1	578	229	206	1	436	189	175	2	366
15-19	612	550	4	1 167	318	357	0	675	291	302	1	594
20-24	1 322	1 263	7	2 591	632	955	3	1 590	587	664	1	1 251
25-29	2 198	2 462	10	4 670	1 222	1 648	6	2 876	982	1 318	3	2 304
30-34	2 936	3 007	15	5 957	1 541	1 751	3	3 296	1 416	1 760	3	3 179
35-39	3 532	2 812	11	6 355	1 684	1 580	3	3 268	1 625	1 680	4	3 310
40-44	3 200	2 537	11	5 748	1 549	1 382	6	2 938	1 589	1 296	1	2 886
45-49	3 262	2 444	4	5 710	1 438	1 153	5	2 596	1 569	1 197	5	2 770
50-54	3 146	2 454	8	5 608	1 289	974	7	2 270	1 398	978	4	2 380
55-59	3 118	2 299	5	5 422	1 275	894	5	2 174	1 474	895	3	2 372
60-64	2 916	2 300	16	5 232	1 138	869	2	2 008	1 304	832	5	2 140
65-69	2 520	1 940	16	4 476	875	710	7	1 592	1 129	783	3	1 916
70-74	2 293	2 196	6	4 495	761	801	2	1 565	1 146	908	5	2 059
75-79	1 781	1 930	9	3 720	553	561	2	1 117	814	974	5	1 793
80-84	1 457	1 783	11	3 251	467	644	2	1 114	730	1 089	4	1 823
85+	1 496	2 716	15	4 227	638	1 143	1	1 782	1 054	2 149	7	3 210
Unspecified	1 064	943	87	2 094	475	404	24	903	389	326	33	748
Total	42 227	38 324	288	80 839	19 147	18 680	117	37 944	20 742	20 033	112	40 887

Table 5: Distribution of household deaths by age, cause of death and sex nationally

Age	Male					Female					RSA				
	Natural	Unnatural	Do not know	Unspecified	Total	Natural	Unnatural	Do not know	Unspecified	Total	Natural	Unnatural	Do not know	Unspecified	Total
0	17 350	1 954	2 011	569	21 885	15 412	1 350	1 620	429	18 812	32 762	3 305	3 631	999	40 697
1-4	5 721	1 110	232	184	7 248	5 189	793	183	139	6 305	10 910	1 904	415	323	13 553
5-9	2 194	740	69	72	3 074	2 112	492	53	63	2 720	4 305	1 232	122	135	5 794
10-14	1 434	651	59	51	2 194	1 420	406	38	29	1 893	2 853	1 057	97	80	4 087
15-19	1 954	2 173	79	72	4 279	2 691	982	78	40	3 792	4 645	3 156	157	113	8 070
20-24	3 964	4 988	146	152	9 251	7 959	1 386	187	103	9 634	11 923	6 374	333	255	18 885
25-29	8 932	5 492	252	263	14 939	14 181	1 644	295	134	16 253	23 113	7 136	547	396	31 192
30-34	13 753	4 277	295	346	18 671	16 006	1 485	311	149	17 950	29 759	5 762	605	494	36 621
35-39	16 212	3 463	312	338	20 324	15 455	1 178	294	162	17 090	31 667	4 641	606	500	37 414
40-44	14 633	2 526	302	287	17 747	12 916	1 034	231	151	14 332	27 549	3 560	533	438	32 079
45-49	14 616	2 376	261	278	17 530	12 065	959	233	141	13 399	26 681	3 335	494	419	30 929
50-54	14 278	1 837	228	283	16 627	11 314	815	199	225	12 552	25 592	2 652	427	508	29 179
55-59	14 257	1 577	200	292	16 326	10 647	650	187	224	11 708	24 903	2 227	387	516	28 034
60-64	14 361	1 134	191	280	15 965	11 140	564	153	216	12 073	25 501	1 698	344	496	28 038
65-69	11 914	751	142	217	13 024	9 645	460	121	183	10 409	21 559	1 211	263	400	23 433
70-74	11 543	590	110	246	12 488	11 157	460	119	209	11 945	22 700	1 049	229	455	24 433
75-79	8 786	363	71	160	9 380	10 086	377	85	187	10 735	18 872	740	156	347	20 115
80-84	7 102	261	61	145	7 568	9 891	289	87	200	10 467	16 993	550	147	345	18 035
85+	7 900	262	63	156	8 381	15 050	370	134	301	15 855	22 950	632	197	457	24 236
Unspecified	4 645	840	135	501	6 121	4 539	380	119	387	5 424	9 184	1 220	254	888	11 546
Total	195 547	37 365	5 219	4 892	243 023	198 875	16 077	4 724	3 672	223 348	394 422	53 441	9 944	8 564	466 371

Note: Total number of deaths excludes 1 696 cases of unspecified sex of the deceased