Cause of death statistics for South Africa: Challenges and possibilities for improvement

Prepared by

Debbie Bradshaw, Victoria Pillay-Van Wyk, Ria Laubscher,
Beatrice Nojilana, Pam Groenewald, Nadine Nannan and Carol Metcalf

Burden of Disease Research Unit

December 2010

Table of Contents

List of Tables

List of Figures

1 Introduction

1.1 Background

Cause of death statistics are an essential data source for monitoring the health of the nations and identifying priorities. They also make an extremely important contribution to the National Burden of Disease Study which attempts to estimate the causes of loss of healthy life.

A review of the quality of cause of death data from 115 WHO member countries, published in 2005, categorised South Africa as having low quality cause of death data based on a combination of a more than 20% of deaths being ill-defined deaths and less than 70% of deaths being registered.[1] Furthermore, extensive miss-classification of HIV/AIDS as an underlying cause,[2,3] made it necessary for the Initial South African National Burden of Disease Study[4] to use a modelling approach to estimate the extent of HIV/AIDS mortality. In addition, the lack of information regarding causes of injury deaths, made it necessary to use the National Injury Mortality Surveillance System to estimate the injury mortality cause profile for the National Burden of Disease Study.

1.2 Validation studies, Cape Town, South Africa

Although there have been major improvements in vital registration in South Africa, and Statistics South Africa considers that by 82% of deaths in 2007 were registered,[5] there is still an increasing concern about the accuracy of the cause of death information for causes other than HIV/AIDS or injuries. Two studies conducted in the Cape Town area have identified relatively high levels of errors made on the death notification forms (DNFs).[6,7] About a third of the DNFs had errors that are serious enough to affect the identification of the underlying cause of death, indicating a clear need for improving cause of death certification.

Further evaluation of the information from the Vangaurd study[6] found that when compared with information obtained from medical records, the cause of death on the DNF agreed for only 50% of the cases at the single NBD cause level.[8] (Medical records were sought from multiple facilities including hospitals, clinics and mortuaries). Correction factors for the cause specific mortality fractions were reported for the more common causes of death (Table 1.1).

While many of the errors resulted from the mis-reporting AIDS as TB and pneumonia on the DNFs, there was a significant number of errors within the cluster of cardiovascular conditions and diabetes. This study also found that out of the 52 cases with ill-defined natural causes (R00-R99) as the underlying cause on the DNF, a specified cause of death could be identified from the medical records in 42 cases accounting (i.e. 81%).

Table 1.1 Cause specific mortality fraction based on the DNF and medical records, (N=703)

Diseases	ICD code	DNF	Medical record	% difference in cause specific mortality fraction	95% CI
HIV	B20-B24, C46	12.0%	18.4%	53.6%	(36.9 to 77.6%)
TB	A15-A19, B90, J90	5.8%	4.3%	-26.8%	(-50.0 to 6.3%)
Pneumonia	J10-J18, J20-J22	2.1%	1.7%	-20.0%	(-60.1 to 59.8%)
Diabetes	E10-E14	6.0%	2.1%	-64.3%	(-77.1 to -37.8%)
Stroke	I60-I69	5.0%	6.5%	31.4%	(-0.7 to 76.6%)
IHD	I20-I25	3.3%	7.3%	121.7%	(53.5 to 228.7%)
Hypertensive diseases	I10-I13	3.3%	5.7%	73.9%	(14.4 to 167.8%)
Chronic obstructive pulmonary disease	J40-J44,I27	3.4%	3.0%	-12.5%	(-42.4 to 31.7%)
Lung cancer	C33-C34	3.7%	3.1%	-15.4%	(-38.7 to 13.8%)
Ill-defined natural	R00 – R99	7.4%	2.1%	-71.2%	(-82.5 to - 3.7%)

Source: Adapted from Burger et al. (submitted)

It not clear whether the errors made on the DNFs observed in Vanguard can be generalised to the whole of South Africa, or that the correction factors can be appropriately applied nationally. Furthermore, the confidence intervals for the corrective factor for single causes based on the Vanguard study are relatively large as a result of the limited sample size. To date, there has been no national validation study to assess the quality of South African cause of death data. This review examines the national cause of death data carefully, as well as the literature reporting on cause of death validation studies, so as to consider the ways in which cause of death data can be improved, assess the need for a national validation study, and identify analytical implications for the planned 2nd National Burden of Disease Study.

2 South African cause of death data, StatSA 2007

Stats SA has made unit record data from 1997-2007 available to the Burden of Disease Research Unit for further analysis, including information about the method of ascertainment of the cause of death as well as the place of occurrence of the death (in hospital, at home etc), and the multiple causes reported on the DNFs. However, it must be noted that in the case of a natural cause of death, it remains possible for a headman to certify a death from natural causes and provide information about the cause of death obtained from the informant using the BI-1680 form.[9] Unfortunately, it is not possible to distinguish the deaths certified by a headman from those certified by a medical practitioner. However we do know that the practice of headman certifying cause of death is generally restricted to rural areas where access to health services is limited.

2.1 Trends in cause of death over a 10 year period

The annual number of registered deaths increased from about 320 000 deaths in 1997 to just over 612 000 in 2006 and dropped slightly in 2007. Figure 2.1 shows that the cause of death profile has changed rapidly during this period with a marked increase in the proportion of deaths due to infectious and parasitic causes and respiratory infections accompanied by a decline in the proportion of deaths due to injuries.

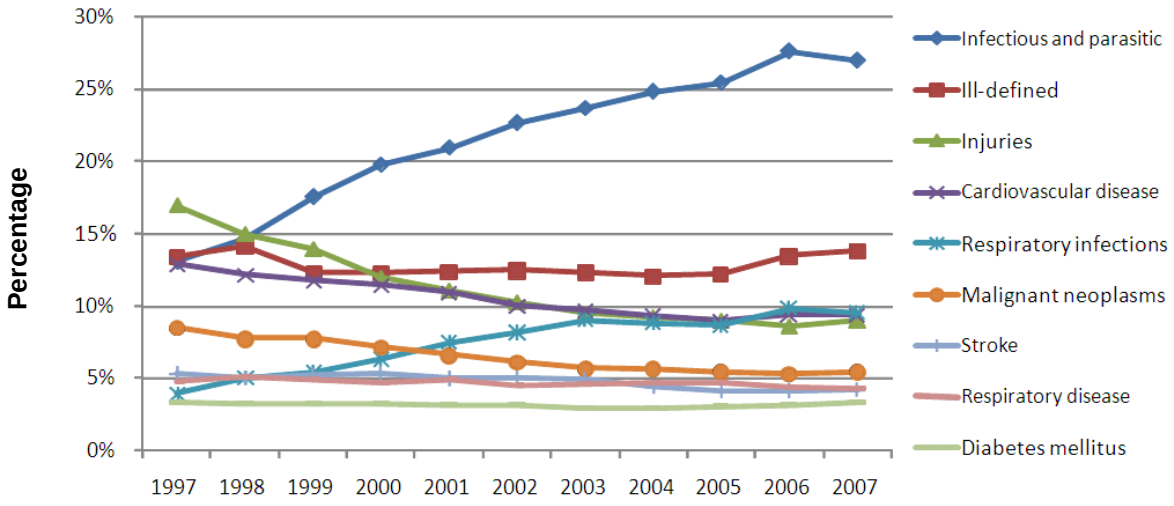


Figure 2.1 The trend in the proportion of deaths due to leading categories, 1997-2007
Source: Statistics South Africa

2.2 Cause of death data, Stats SA 2007

2.2.1 Ill defined deaths and garbage codes, 2007

The cause of death information for 2007 has been examined according to the new NBD list (N=601 144). Selected causes based on their common frequency, lack of detail about the actual underlying cause (e.g. epilepsy), or evidence of certification errors (e.g. diabetes) are listed in a summary of the deaths in 2007 in Table 2.1. The relative frequency is reported as well as the average number of lines of information on the DNF for that cause. The average number of lines of information on the DNF broadly indicates the level of detail provided on the DNF. The last 2 columns show the proportion of deaths that occurred in hospital. Since there was considerable missing information (about 20% did not have place of death specified), the proportion is calculated out of the total as well as the total that had completed information specified.

Table 2.1 Summary of the 2007 cause of death data, Stats SA

Cause	Number		%	Average lines on DNF	% in hospital	
					Total	Specified
Injuries (including undetermined and garbage codes)	54147		9.0%	**1.16**	**22.3%**	**30.4%**
Specified natural causes	409751		68.2%	**1.71**	**53.3%**	**59.9%**
HIV		38751	6.4%	2.25	68.5%	76.2%
Diabetes		20139	3.4%	2.33	50.0%	55.7%
Epilepsy		3660	0.6%	1.63	35.0%	40.5%
Other specified		347200	57.8%	1.62	52.0%	58.5%
Ill-defined natural	**83210**		**13.8%**	**1.16**	**7.5%**	**10.6%**
Garbage code (natural)	**53966**		**9.0%**	**1.50**	**50.3%**	**57.2%**
Septicaemia		5186	0.9%	1.57	68.9%	76.5%
Volume depletion and other disorder of fluid		4351	0.7%	1.75	58.7%	64.9%
Hepatic failure, not elsewhere classified		1964	0.3%	1.52	71.0%	78.2%
Renal failure		6523	1.1%	1.64	70.3%	78.4%
Heart failure		16465	2.7%	1.29	34.7%	39.3%
Haematemesis & melaena/other diseases of GIT		1069	0.2%	152	58.9%	69.5%
Other garbage		18407	3.1%	154	47.3%	55.2%
Total	**601144**		**100.0%**	**1.57**	**43.9%**	**51.6%**

Source: Statistics South Africa

From Table 2.1, it can be seen that there were 13.8% deaths due to ill-defined natural causes and the vast majority of these occurred outside of hospital. A further 9.0% of deaths were due to natural 'garbage' codes, half of which were outside hospital. 'Garbage' codes, is a category used in the Global Burden of Disease Study for those codes which do not signify an underlying cause of death, including intermediate, immediate, and unlikely or ambiguous causes of death.[10]

In some of the rural areas in SA it is common practice for headman (Appendix) to certify cause of death. However, the death data does not have a field indicating whether the death certified by a doctor or a headman. Using the absence of a SAMDC or SANC registration number to indicate the cases certified by headman, we found that ±10% of deaths in 2007 were certified by headman (Dr Maletela Tuoane-Nkhasi, StatsSA – personal communication). Half of these had information about the cause of death from an informant, allowing a specified cause to be coded. The other 5% were classified as ill defined (Figure 2.2).

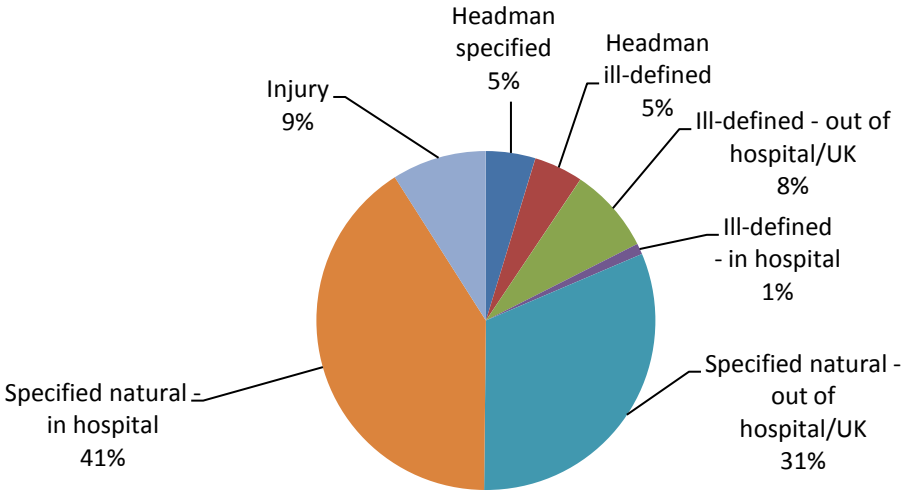


Figure 2.2 Cause of death certification in 2007
Special tabulation provided by Dr Maletela Tuoane-Nkhasi, StatsSA

2.2.2 Leading specific causes of death, 2007

Table 2.2 shows the top 25 causes of death in 2007 which account for just over 80% of all the deaths. It should be noted that tuberculosis has been sub-divided into respiratory, extra-pulmonary and military disease, each of which appears among the leading causes. When combined, tuberculosis accounted for 12.9% of the total deaths. Two of the garbage codes namely heart failure (2.7%) and septicaemia (0.9%) are also listed in of the top 25 causes of death.

Table 2.2 Top 25 causes of death in 2007, Stats SA

Cause	N	Total
Ill defined natural	83,208	13.8%
Respiratory tuberculosis	66,500	11.1%
Lower respiratory infections	58,375	9.7%
Intestinal infectious diseases	40,850	6.8%
HIV/AIDS	38,751	6.4%
Cerebrovascular disease	25,321	4.2%

Diabetes mellitus	20,139	3.4%
Injuries with undetermined intent	20,128	3.3%
Heart failure	16,465	2.7%
Hypertensive heart disease	13,381	2.2%
Ischaemic heart disease	12,506	2.1%
Other infectious diseases	8,981	1.5%
Exposure to unspecified factor	8,930	1.5%
COPD	8,867	1.5%
Meningitis	8,427	1.4%
Endocrine nutritional, blood and immune disorders	6,305	1.0%
Asthma	6,228	1.0%
Road injuries	6,038	1.0%
Extrapulmonary tuberculosis	5,749	1.0%
Septicaemia	5,186	0.9%
Miliary tuberculosis	5,008	0.8%
Mechanical forces; firearm	4,903	0.8%
Interpersonal violence without firearm	4,813	0.8%
Trachea/bronchi/lung cancer	4,689	0.8%
Other respiratory disease	4,386	0.7%
Top 25 causes	484,134	80.5%
All causes	601,133	100.0%

Source: Statistics South Africa

2.2.3 Place of death for leading causes of death, 2007

Overall, 43.9% of deaths occurred in hospital. Table 2.3 shows the place of death of the leading 25 causes of death. The proportion of deaths that occurred at outpatient wards was small (1.8%) and has been combined with the deaths that occurred in hospital. The home deaths have been combined with a small proportion of deaths in nursing homes (2.1%) and added to the other category (4.4%). Overall 15% of cases had missing information about the place of death. While some forms would have this field missing, it is possible that much of the missing data would be for deaths certified by a headman, which does not have a field for this detail. The injury deaths also have a relatively high proportion with missing information. Only 7.5% of the deaths due to ill-defined natural causes occurred in hospital.

Table 2.3 Place of death for top 25 causes in 2007, Stats SA

Cause	Hospital/ Outpatient	Other	Unknown	Total	N
Ill defined natural	7.5%	63.2%	29.3%	100.0%	83,208
Respiratory tuberculosis	58.4%	31.5%	10.1%	100.0%	66,500
Lower respiratory infections	43.4%	44.9%	11.7%	100.0%	58,375
Intestinal infectious diseases	48.3%	41.1%	10.7%	100.0%	40,850
HIV/AIDS	68.5%	21.4%	10.1%	100.0%	38,751
Cerebrovascular disease	46.9%	41.0%	12.1%	100.0%	25,321
Diabetes mellitus	50.0%	39.7%	10.4%	100.0%	20,139
Injuries with undetermined intent	21.0%	51.5%	27.5%	100.0%	20,128
Heart failure	34.7%	53.8%	11.5%	100.0%	16,465
Hypertensive heart disease	27.4%	60.1%	12.5%	100.0%	13,381
Ischaemic heart disease	29.8%	58.0%	12.2%	100.0%	12,506

Other infectious diseases	78.7%	13.8%	7.5%	100.0%	8,981
Exposure to unspecified factor	28.6%	48.6%	22.8%	100.0%	8,930
COPD	41.3%	47.7%	11.0%	100.0%	8,867
Meningitis	76.1%	15.3%	8.5%	100.0%	8,427
Endocrine nutritional, blood and immune disorders	71.1%	20.0%	8.9%	100.0%	6,305
Asthma	17.3%	63.6%	19.1%	100.0%	6,228
Road injuries	18.7%	53.5%	27.8%	100.0%	6,038
Extrapulmonary tuberculosis	82.1%	9.7%	8.2%	100.0%	5,749
Septicaemia	68.9%	21.2%	9.9%	100.0%	5,186
Miliary tuberculosis	83.5%	6.9%	9.6%	100.0%	5,008
Mechanical forces; firearm	15.8%	53.7%	30.5%	100.0%	4,903
Interpersonal violence without firearm	13.8%	54.3%	31.8%	100.0%	4,813
Trachea/bronchi/lung cancer	49.0%	40.0%	11.0%	100.0%	4,689
Other respiratory	34.2%	43.6%	22.1%	100.0%	4,386
Top 25 causes	41.3%	43.0%	15.6%	100.0%	484,134
All causes	43.9%	41.1%	15.0%	100.0%	601,133

Source: Statistics South Africa

2.2.4 Natural deaths due to ill-defined causes, 2007

Ill-defined causes accounted for 15.2% of natural causes but differed across age groups (Figure 2.3). Over 20% of the natural deaths in the elderly (80+years old) were due to ill-defined causes, considerably higher than the national average of 15.2%. This could be a result of multiple conditions associated with old age making it difficult to diagnose the underlying cause of death, or the death occurring out of hospital more frequently. The proportion natural deaths due to ill-defined causes for children aged 1-14 years was also higher than the average.

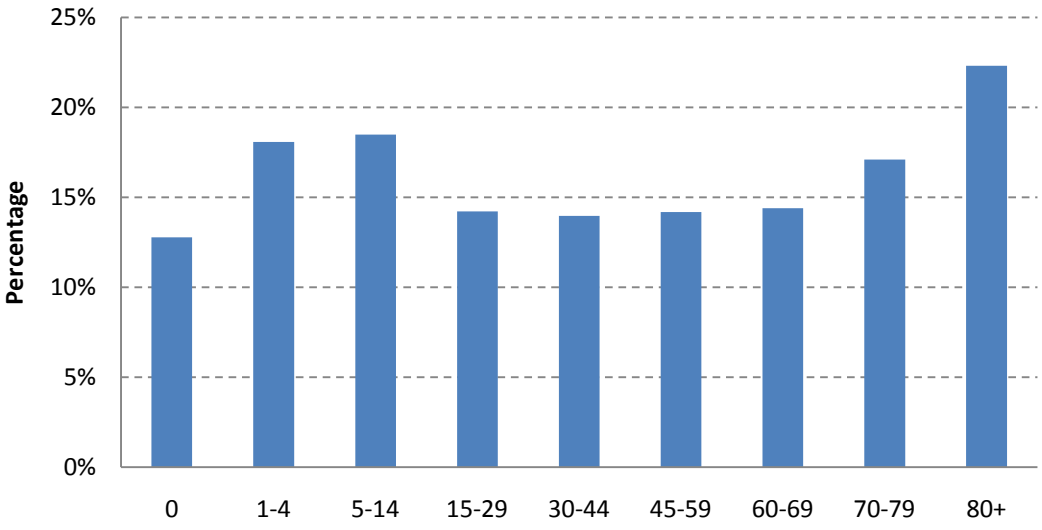


Figure 2.3 Proportion of natural deaths due to ill define natural causes by age group, 2007

Source: Statistics South Africa

The Western Cape Burden of Disease Study currently underway identified a higher proportion of infant deaths due to ill-defined causes (Dr P Groenewald, MRC – personal communication) and prompted further analysis of ill-defined cause of death for infants stratified by province. Table 2.4 shows that overall 12.4% of infant deaths are ill-defined - with Limpopo province having the highest proportion (16.3%), followed by the Western Cape (15.5%). All the provinces had similarly low proportions of the ill-defined infant deaths that occurred in hospital (5.8%) – data not shown. Nearly a quarter (24.1%) of the ill-defined infant deaths in the Western Cape were based on autopsy compared with 3.2% nationally. Of the ill- defined infant deaths, half (49.9%) in the Western Cape had an autopsy compared to the national average of 5.4%. A possible interpretation for this striking observation is that the Western Cape has fewer deaths from the easily diagnosed causes (e.g. diarrhoea, pneumonia etc) and hence, despite better access to forensic services, the proportion of ill-defined deaths are higher in this province.

Table 2.4 Proportion ill defined infant deaths by province and autopsy ascertainment

Province	Total deaths	Ill-defined		Autopsy		Ill-defined with autopsy	
	N	n	% of total	n	% of total	n	% with autopsy
Western Cape	2,646	411	15.5%	637	24.1%	205	49.9%
Eastern Cape	4,099	619	15.1%	136	3.3%	18	2.9%
Northern Cape	1,230	189	15.4%	94	7.6%	15	7.9%
Free State	4,733	469	9.9%	97	2.0%	6	1.3%
Kwa-Zulu Natal	9,877	1,017	10.3%	148	1.5%	23	2.3%
North West	4,966	582	11.7%	55	1.1%	5	0.9%
Gauteng	10,831	1,569	14.5%	228	2.1%	29	1.8%
Mpumalanga	3,933	241	6.1%	46	1.2%	3	1.2%
Limpopo	4,236	692	16.3%	56	1.3%	10	1.4%
Foreign	2	1	50.0%	0	0.0%	0	0.0%
Total	**46,553**	**5,790**	**12.4%**	**1,497**	**3.2%**	**314**	**5.4%**

Source: Statistics South Africa

2.2.5 Age specific death rates for natural causes and external causes, 1997-2007

The trend in the age specific death rates for natural causes and for the external causes separately are shown in Figure 2.4. While deaths from natural causes increased markedly in the young adult ages and children, there was also some increase in the older ages over this period. In contrast, death rates from external causes declined over this period, particularly in the earlier period. The age pattern remained fairly consistent with the exception of older ages, possibly reflecting some age mis-classification in these data.

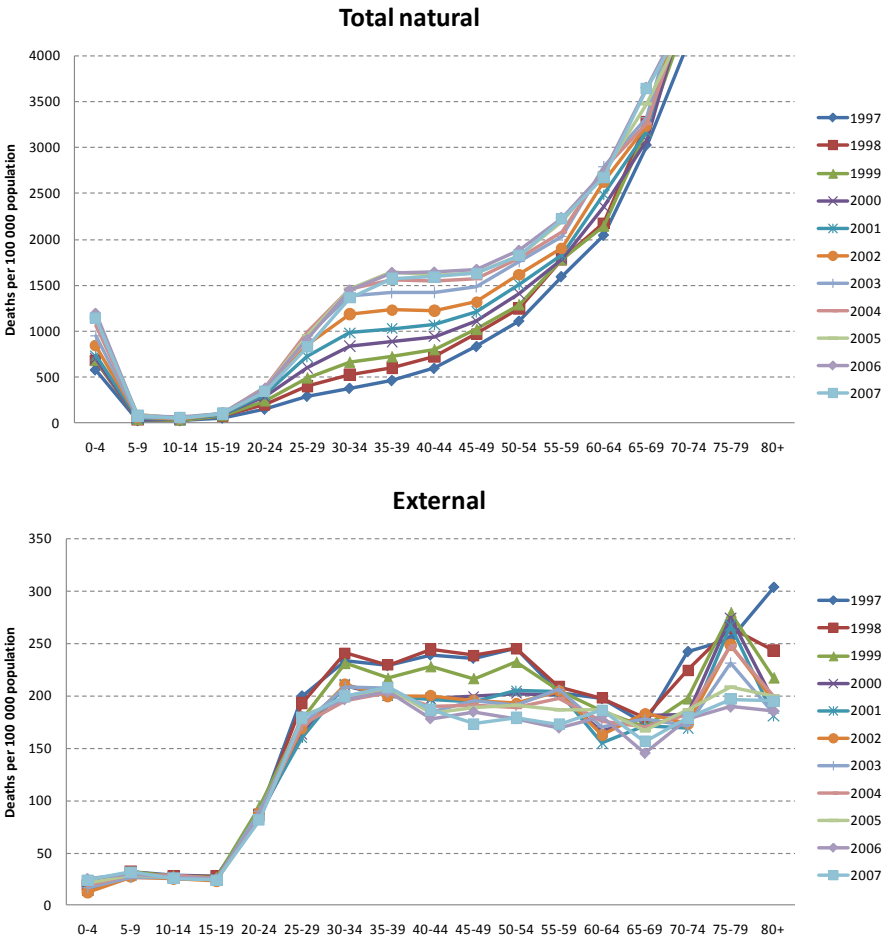


Figure 2.4 Age specific rates for natural and external causes of deaths by age group, 1997-2007
Source: Statistics South Africa and ASSA2003

2.2.6 Age specific death rates for commonly reported infection diseases, 1997-2007

Previous analysis of the cause specific deaths rates has provided valuable insight into the mis-classification of AIDS deaths. The age distributions are presented for the most common causes, as well as the identified garbage codes. Figure 2.5 shows the trend in the rates for the commonly reported infectious diseases such as respiratory TB, intestinal infections etc. These infectious diseases display the trend seen in the HIV/AIDS death rates i.e. a rapid increase in young adult ages, as well as children, suggesting that HIV is often not reported as the underlying cause of death but rather the indicator condition. Some of the causes have higher deaths rates in the older ages and show no significant increase over the 10 year period. It is interesting to see an increase in HIV/AIDS over the age of 80 years (about 150 cases per year), and the increase in intestinal infections, meningitis and other infectious diseases.

2.2.7 Age specific death rates for commonly reported non-communicable diseases, 1997-2007

Figure 2.6 shows the age specific death rates for the commonly reported non-communicable diseases such as cerebrovascular disease (stroke), diabetes, ischaemic heart disease and hypertensive disease. These all reflect the typical increase over the age groups with year on year increases in diabetes, hypertensive heart disease, asthma, epilepsy and other endocrine. There was a decrease in chronic obstructive pulmonary disease and ischaemic heart disease over the age of 70 years. The HIV age pattern does not appear in any of the common conditions with the exception of the endocrine nutritional and other blood immune disorders to which “immune suppression” would be coded. The increase in epilepsy indicates that some of the increase may be related to HIV but that there has also been an increase across the older ages. Other factors may account for some of or all of the observed increase.

2.2.8 Age specific death rates for commonly reported garbage codes for natural causes, 1997-2007

Figure 2.7 shows the trend in the age specific death rates for the common garbage codes within the natural causes. The age pattern for heart failure suggests that this cause of deaths has not been influenced by HIV. However, ill-defined causes and volume depletion include a definite HIV pattern. Ill defined deaths were adjusted by removing misclassified HIV using the proportion of ill defined deaths reported in 1997 as an estimate for baseline (no HIV deaths). After the adjustment, ill defined deaths dropped by 3.7% i.e. 13.8% to 10.1%. A similar decrease was noted for ill defined deaths among males (12.8% to 9.2%) and females (15.0% to 11.1%). Septicaemia and renal failure have small peaks in the young adult ages that could be associated with HIV trends.

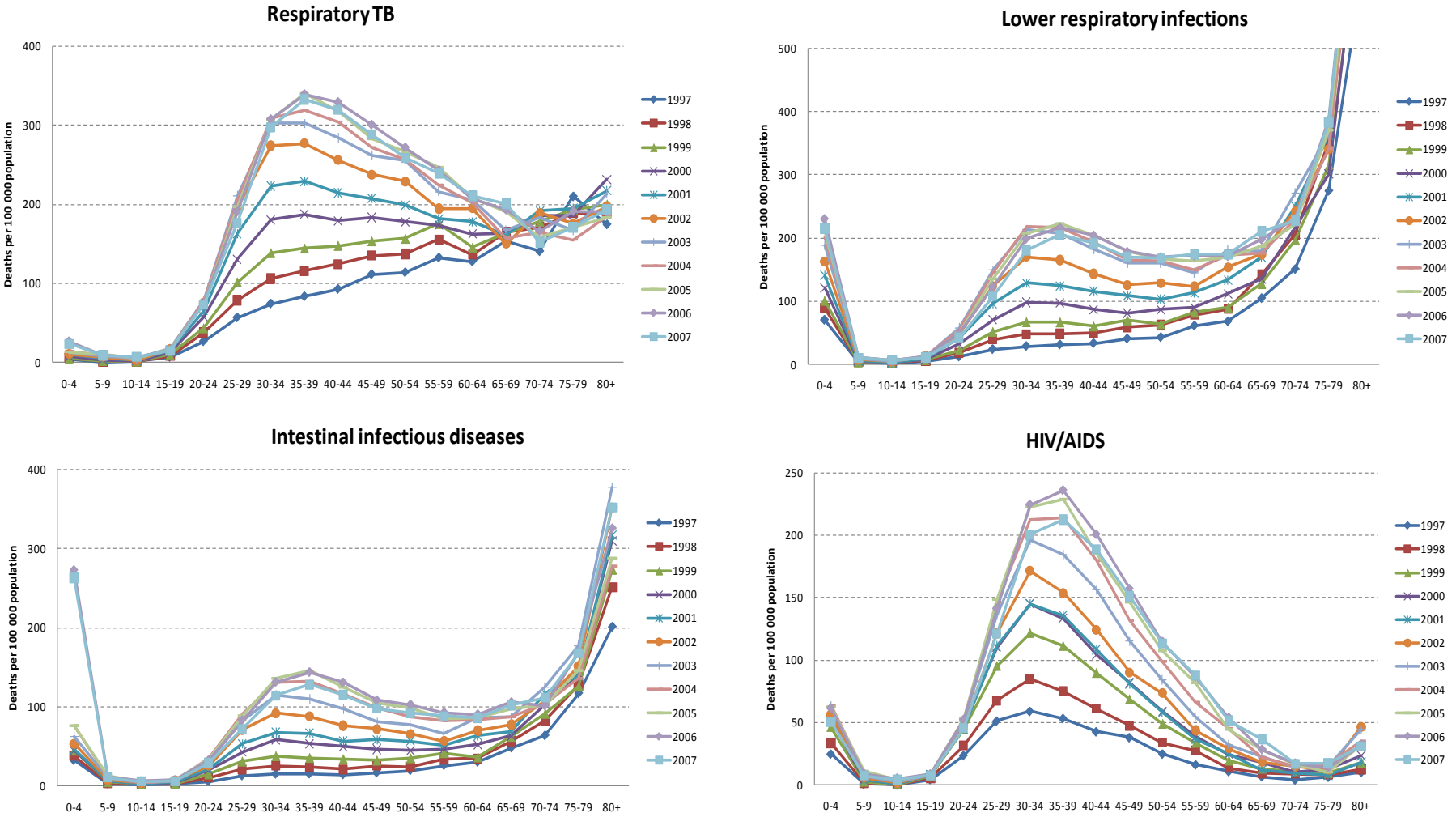


Figure 2.5 Age specific rates for common infectious disease by age group, 1997-2007

(continued on next page)

Source: Statistics South Africa and ASSA2003

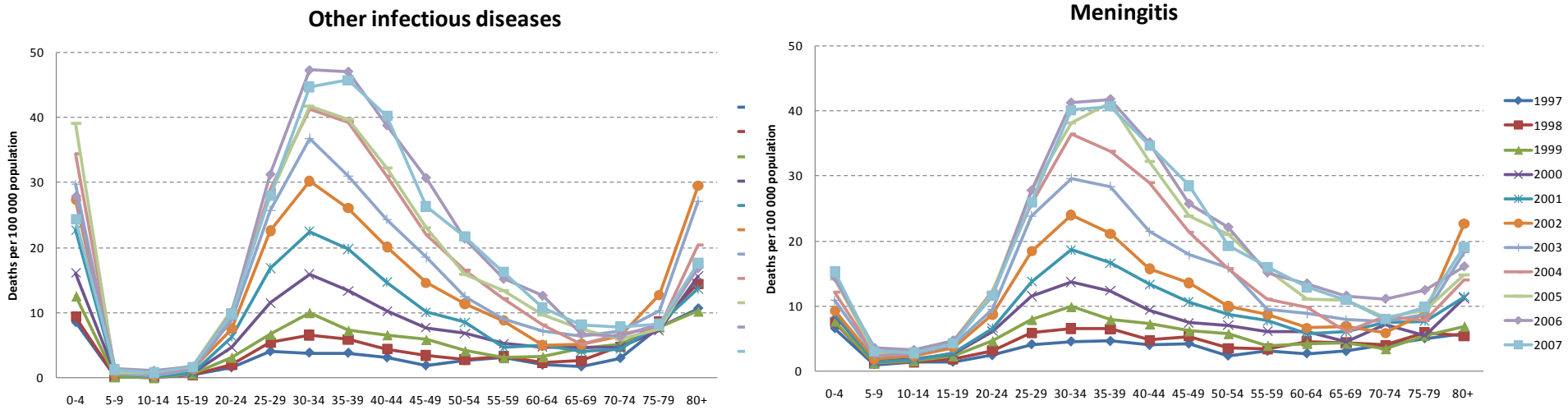


Figure 2.5 Age specific rates for common infectious disease by age group, 1997-2007
(continued from previous page)
Source: Statistics South Africa and ASSA2003

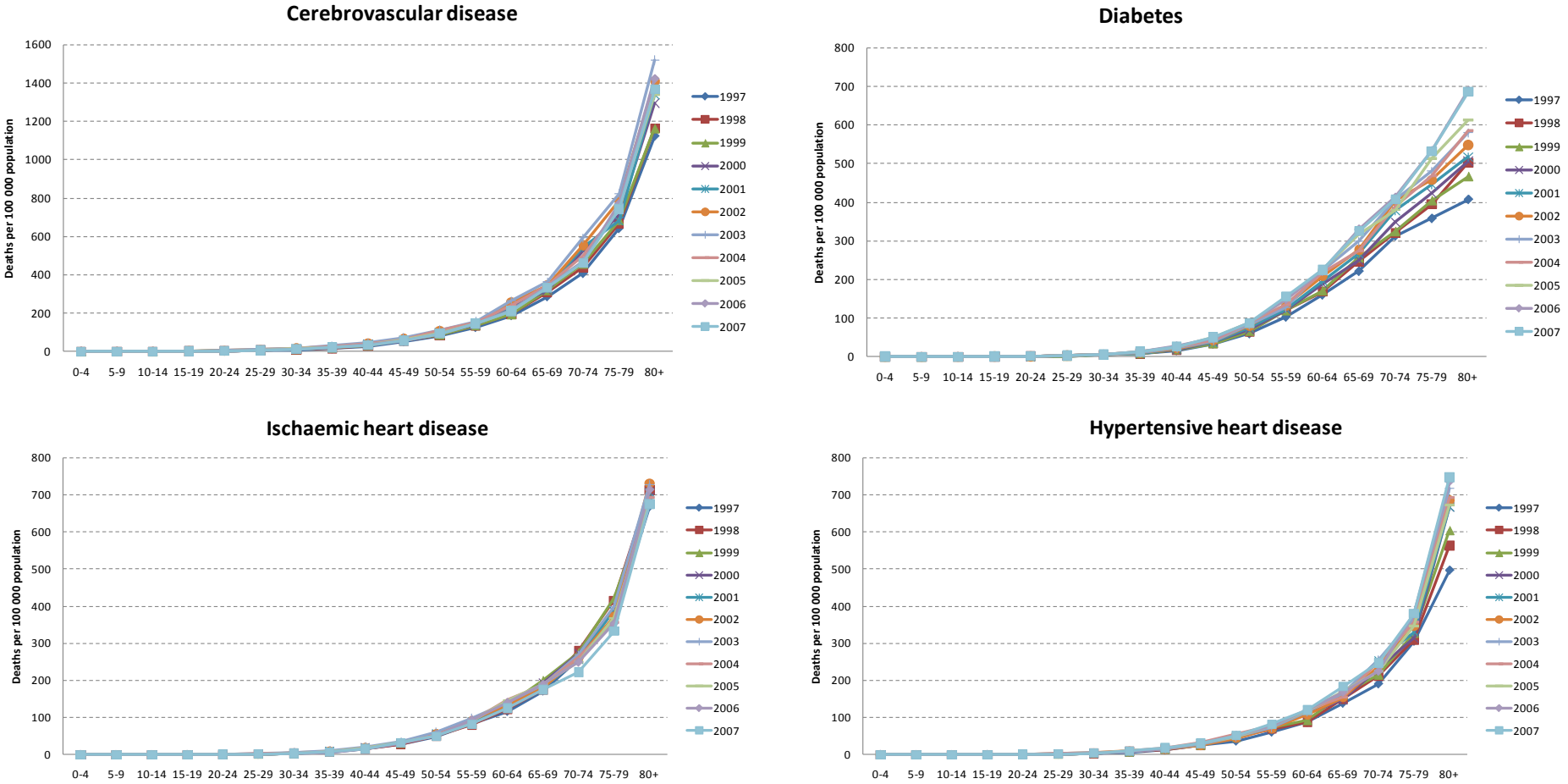


Figure 2.6 Age specific rates for common non-communicable diseases by age group, 1997-2007

(continued on next page)

Source: Statistics South Africa and ASSA 2002

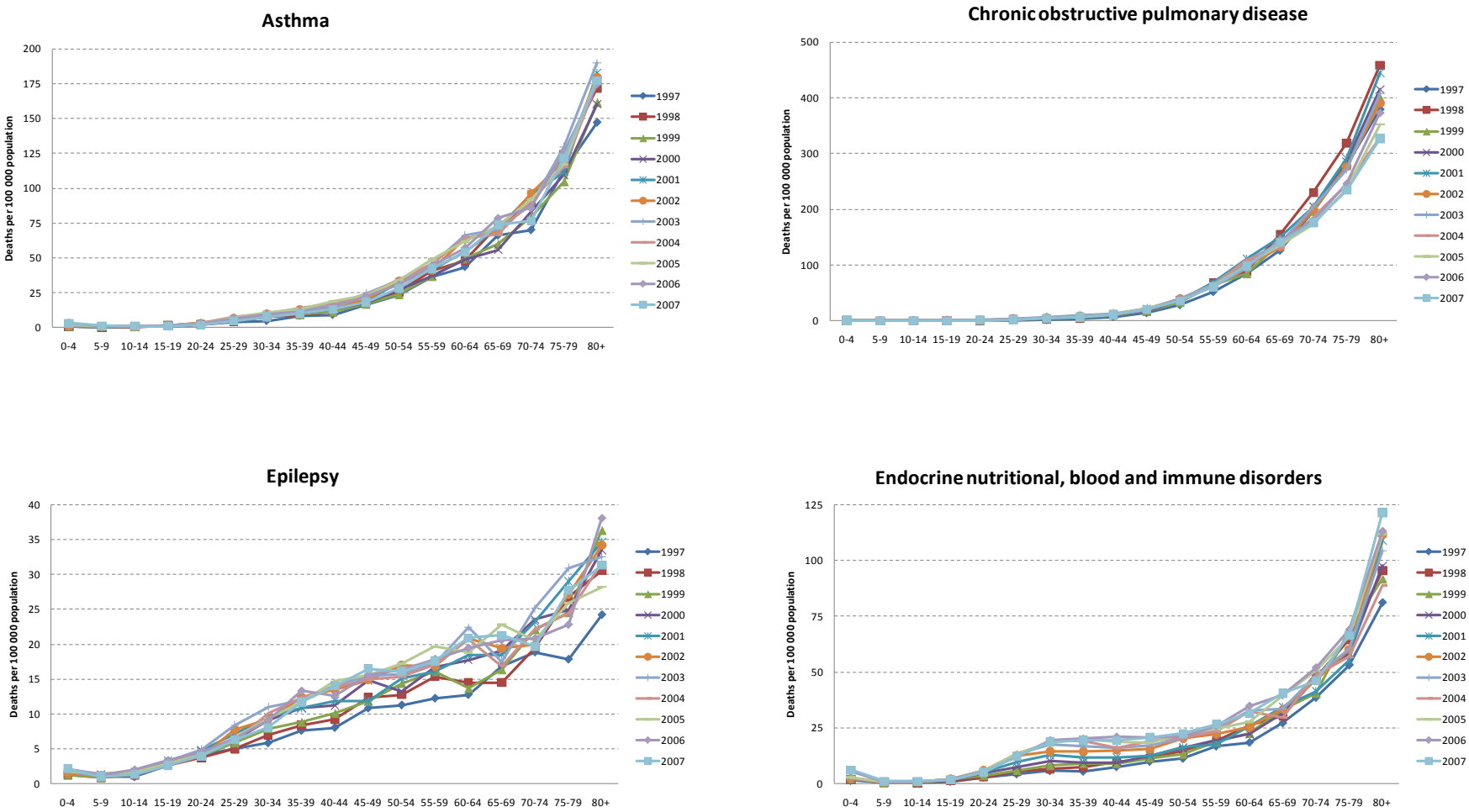


Figure 2.6 Age specific rates for common non-communicable diseases by age group, 1997-2007
(continued from previous page)
Source: Statistics South Africa and ASSA 2002

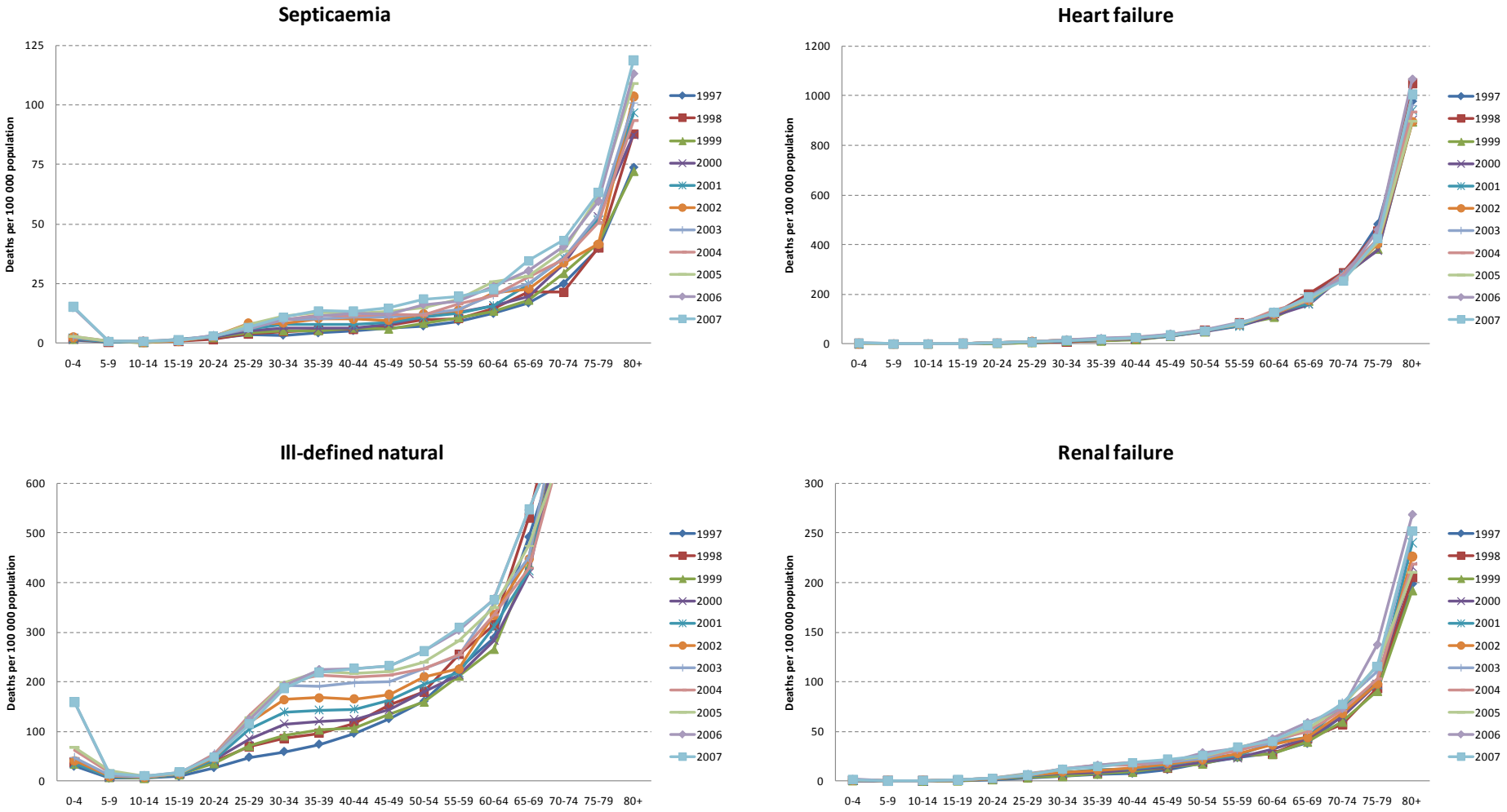


Figure 2.7 Age specific rates for common garbage codes (natural) by age group, 1997-2007

(continued on next page)

Source: Statistics South Africa and ASSA2003

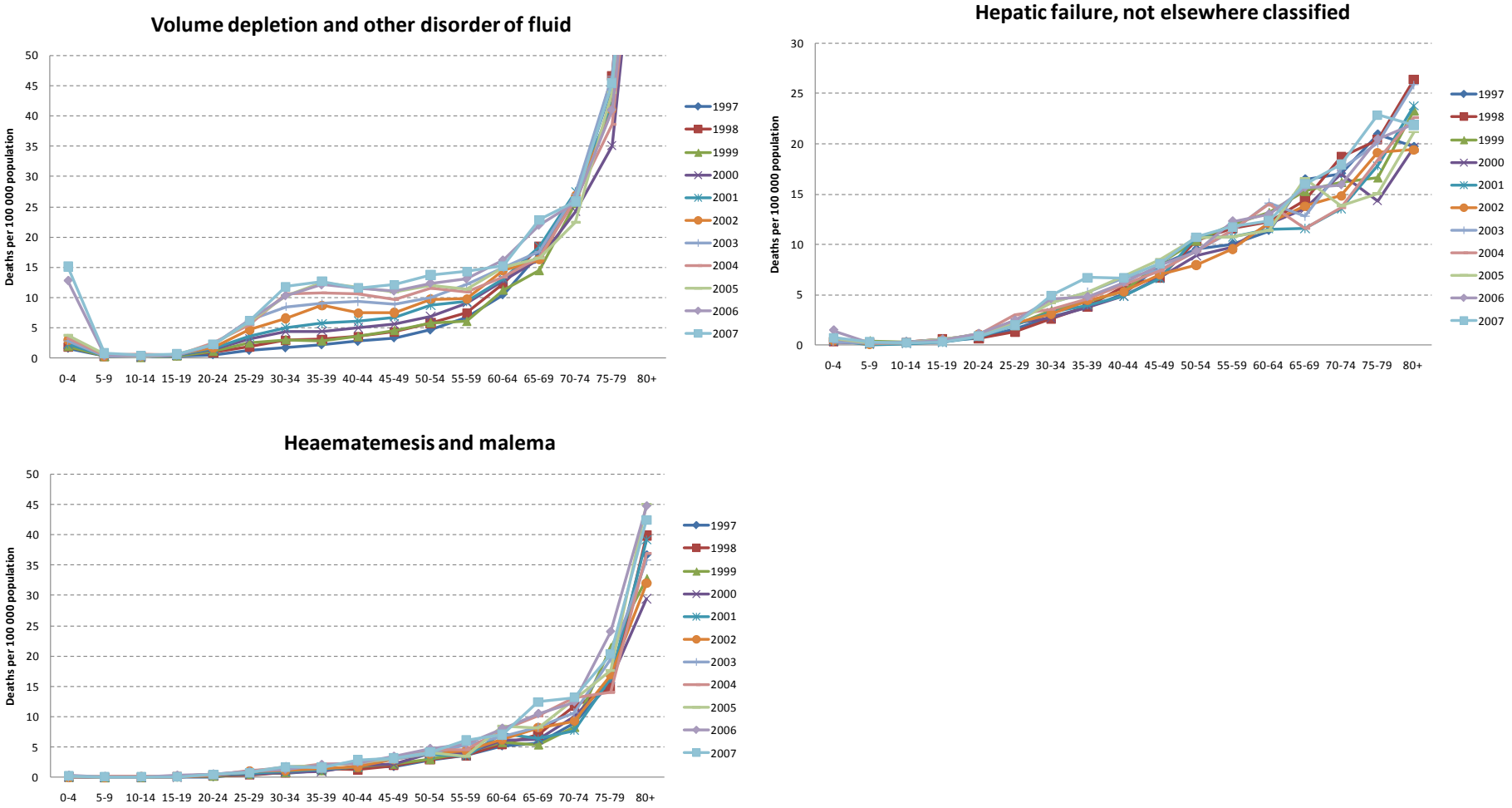


Figure 2.7 Age specific rates for common garbage codes (natural) by age group, 1997-2007
(continued from previous page)
Source: Statistics South Africa and ASSA2003

2.2.9 Age specific death rates for external cause garbage codes, 1997-2007

Figure 2.8 shows the trend in age specific rates for two of the external cause garbage codes i.e. injuries with undetermined intent and exposure to unspecified factor. In 1997, the undetermined intent accounted for 84% of the external causes and account for the majority of injury deaths. With the introduction of the new death certificate in 1998, it would appear that there was a drop in injuries with undetermined intent to 70% and this remained at about two thirds until 2006. However, in 2007 the undetermined was reduced to 37%. In Figure 2.8, it can be seen that the rates were halved – suggesting that the quality of the injury cause of death in vital statistics may have improved. However, the deaths from accidental exposure to unspecified factors, still not an adequately specified cause of death, increased in 2007.

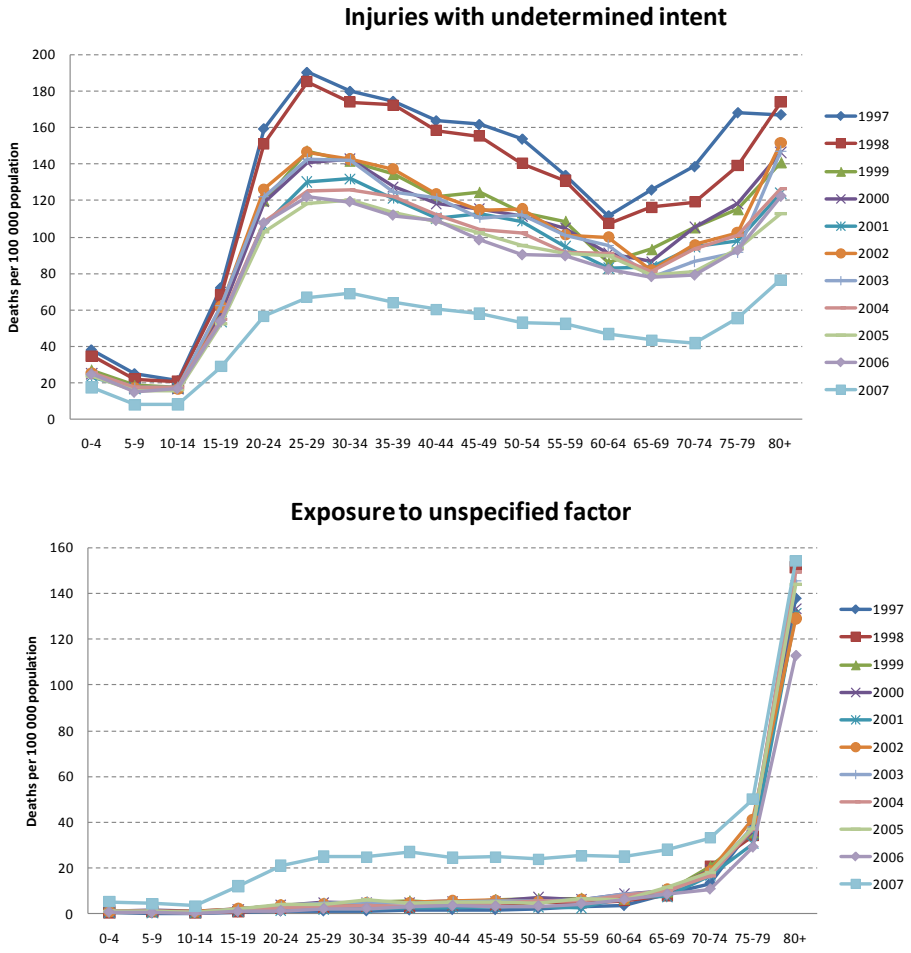


Figure 2.8 Age specific rates for common garbage codes (external) by age group, 1997-2007

Source: Statistics South Africa and ASSA2003

2.2.10 Multiple cause of death information

The multiple cause information has been analysed for selected causes. Diabetes has been included because of the errors identified in certification of diabetes as a cause.[7,8] In addition, multiple cause information has been analysed for heart failure and epilepsy in order to assess whether the multiple cause information provides additional insight into the causes.

Diabetes

The average number of lines completed on the DNFs which have diabetes as an underlying cause went from an average of 2.6 in 1997 to 2.3 in 2007. In 2007, there were 46912 multiple causes of death listed in Part I and Part II for the cases with diabetes as an underlying cause (Table 2.5). Of these causes, the majority were diabetes (43.1%). Thereafter, cardiovascular conditions were the most common multiple cause accounting for 33.1%. It should be noted that acute or terminal circulatory conditions which, according to the ICD-10 coding rules, can be included in Part I when diabetes is the underlying cause of death, accounted for 14.5% of the multiple causes. Approximately 6.2% were mechanisms and 2.1% were intermediate conditions that could have followed from diabetes.

Table 2.5 Top 20 multiple causes of death when diabetes is underlying cause, 2007

	Part I & II		Part I		Part II	
Cause	**N**	**% of total**	**N**	**% of total**	**N**	**% of total**
Diabetes Mellitus	20,238	43.1	20,119	43.4	119	23.2
Hypertensive disease	7,599	16.2	7,535	16.2	64	12.5
Cerebrovascular disease	3,176	6.8	3,154	6.8	22	4.3
Ischaemic heart disease	2,341	5.0	2,323	5.0	32	6.2
Lower respiratory infections	1,421	3.0	1,409	3.0	18	3.5
Septicaemia	1,456	3.1	1,444	3.1	12	2.3
Heart failure	1,060	2.3	1,059	2.3	1	0.2
Cardiac arrest	850	1.8	848	1.8	-	-
Unspecified renal failure	835	1.8	823	1.8	12	2.3
Ill-defined natural	696	1.5	687	1.5	9	1.8
Endocrine nutritional, blood and immune disorders	868	1.9	836	1.8	32	6.2
Chronic renal failure	514	1.1	508	1.1	6	1.2
Gangrene, not elsewhere classified	354	0.8	351	0.8	3	0.6
Skin diseases	312	0.7	303	0.7	9	1.8
Other disorders of fluid, electrolyte and acid-base balance	383	0.8	378	0.8	2	0.4
Acute renal failure	245	0.5	244	0.5	-	-
Volume depletion	189	0.4	184	0.4	5	1.0
Other urinary and gynaecological diseases	213	0.5	207	0.5	4	0.7
Pulmonary oedema	208	0.4	209	0.4	-	-
Respiratory failure, not elsewhere classified	173	0.4	171	0.4	-	-
Top 20 causes	**43,131**	**91.9**	**42,412**	**92.2**	**350**	**68.1**
Total multiple causes	**46,912**	**100.0**	**46,398**	**100.0**	**514**	**100.0**

Source: Statistics South Africa

Table 2.6 shows the proportion of diabetes deaths that also had cardiovascular diseases reported in Part I. The overall proportion was 54.8% for the period 1997-2007. The proportion of hypertensive diseases increased from about 25% to 37% and was more or less double the proportion with stroke. Stroke occurred in 15.6% and ischaemic heart disease in 10.9% over with little change over the period.

A quarter (26.7%) of the deaths had essential hypertension also reported in Part 1 when diabetes was reported in Part 1. The same was observed for essential hypertension heart disease (6.3%), essential hypertension renal disease (3.2%) and hypertension heart and renal disease (0.6%).

Table 2.6 The proportion of diabetes deaths with cardiovascular diseases recorded in Part I, 1997-2007

Year	Diabetes	% with cardiovascular disease	% with hypertensive disease	% with Stroke	% with IHD
1997	11,705	57.7%	25.6%	15.3%	11.0%
1998	13,581	57.2%	28.2%	16.0%	10.7%
1999	13,746	56.1%	30.2%	16.1%	12.2%
2000	14,973	55.6%	32.2%	16.5%	11.9%
2001	15,879	54.8%	32.1%	17.0%	11.9%
2002	17,260	54.5%	33.4%	16.3%	11.1%
2003	18,331	54.6%	34.7%	15.8%	10.9%
2004	18,488	53.9%	35.2%	15.2%	10.7%
2005	20,178	53.6%	35.7%	15.4%	10.3%
2006	21,445	53.8%	36.4%	14.6%	10.0%
2007	21,915	53.6%	36.8%	14.6%	9.9%
Total	**187,501**	**54.8%**	**33.4%**	**15.6%**	**10.9%**

Source: Statistics South Africa

Heart failure

On average the DNF for the cases with heart failure as the underlying cause of death had 1.39 lines completed resulting in 21519 multiple causes. The causes ranged from ill-defined causes to specific conditions (Table 2.7). Heart failure accounted for the large majority of causes (77.4%). However of all the others, the majority were ill-defined or mechanisms which accounted for 72.9%. Specific cardiovascular conditions (cerebrovascular disease, other circulatory disease, conduction disorders etc) accounted for 9.3% and other specified causes (e.g. COPD, septicaemia etc) accounted for 17.8%. Of the 16465 cases with heart failure as the underlying cause, just under 40% of the cases that had heart failure as the underlying cause of death occurred in hospital and 52% were ascertained by a medical doctor or autopsy.

Table 2.7 Top 20 multiple causes when heart failure is the underlying cause

Multiple cause	N	% of total	% of other
Heart failure	16,444	77.4%	-
Ill-defined natural	827	3.9%	17.2%
Lower respiratory infections	629	3.0%	13.1%
Cardiac arrest	561	2.6%	11.7%
Pulmonary oedema	505	2.4%	10.5%
COPD	265	1.2%	5.5%
Cerebrovascular disease	179	0.8%	3.7%
Respiratory failure, not elsewhere classified	177	0.8%	3.7%
Complications and ill-defined descriptions of heart disease	149	0.7%	3.1%
Other circulatory diseases	133	0.6%	2.8%
Unspecified renal failure	126	0.6%	2.6%
Septicaemia	124	0.6%	2.6%
Ill defined: redistribute on some causes	101	0.5%	2.1%
Asthma	98	0.5%	2.0%
Endocrine nutritional, blood and immune disorders	74	0.3%	1.5%
Other muculo-skeletal	65	0.3%	1.4%
Pulmonary embolism	56	0.3%	1.2%
Acute renal failure	55	0.3%	1.1%
Chronic renal failure	55	0.3%	1.1%
Conduction disorders and other dysrythmias	54	0.3%	1.1%
Top 20 causes	**20,677**	**97.3%**	**88.1%**
Total multiple causes	**21,251**	**100.0%**	**100.0%**

Source: Statistics South Africa

Epilepsy

On average the DNF for the cases with epilepsy as the underlying cause of death had 1.63 lines completed resulting in 5977multiple causes. Very few causes were reported in part 2 (28 cases). Epilepsy accounted about two thirds of the causes (63.9%) (Table 2.8). Cerebrovascular disease was the most common multiple cause for cases with epilepsy as the underlying cause, and accounted for 12% of the multiple causes. The majority of the cases were recorded in Part I of the DNF, together with epilepsy. It is strange that such cases did not have cerebrovascular disease identified as the underlying cause of death and should be investigated further.

Table 2.8 Top 20 multiple causes when Epilepsy is the underlying cause

Multiple cause	N	% of total	% of other
Epilepsy	3,803	63.9%	-
Cerebrovascular disease	260	4.4%	12.0%
Ill defined natural: redistribute all causes	228	3.8%	10.5%
Lower Respiratory infections	205	3.4%	9.4%
Pneumonitis due to solids and liquids	136	2.3%	6.3%
Hypertensive heart disease	131	2.2%	6.0%
Cardiac arrest	87	1.5%	4.0%
Heart failure	87	1.5%	4.0%
Other disorders of brain (encephalopathy)	77	1.3%	3.5%
Septicaemia	55	0.9%	2.5%
Respiratory failure, not elsewhere classified	50	0.8%	2.3%
Diabetes Mellitus	47	0.8%	2.2%
Respiratory Tuberculosis	41	0.7%	1.9%
Schizophrenia	41	0.7%	1.9%
Other threats to breathing	40	0.7%	1.8%
HIV/AIDS	39	0.7%	1.8%
Mental Retardation not included as sequelae elsewhere	35	0.6%	1.6%
Meningitis	31	0.5%	1.4%
Other respiratory	31	0.5%	1.4%
Asthma	26	0.4%	1.2%
Top 20 causes	**5,450**	**91.2%**	**75.8%**
Total multiple causes	**5,977**	**100.0%**	**100.0%**

Source: Statistics South Africa

3 Literature review on cause of death validation studies

3.1 Purpose for review

In light of the anomalies noted in the South African cause of death data, a literature review was conducted to identify studies that have been done elsewhere to resolve the underlying cause of death, and more specifically to look at the methodological approaches used that could possibly inform the design of a validation study for South Africa.

3.2 Literature review of cause of death validation studies, 2000-2010

3.2.1 Published review identified

A 2006 review performed by Johansson *et al.*[11] for the period 1998 – 2004 identified 44 cause of death validation studies (12 on all causes of death, and 32 on selected conditions or groups of conditions) and provided a critique of the overall quality of cause of death validation studies. This review highlighted that the methods used were often not described in sufficient detail, that often the studies did not report criteria for identifying the underlying cause of death when several independent conditions were present, and that in several studies the criteria used to select the main cause of death when faced with competing causes of death conflicted with international WHO standards. They also found that in many studies the researchers failed to draw a distinction between the principal underlying cause of death and contributory causes of death, or to specify the criteria for selecting the principal underlying cause of death. Johansson *et al.*[11] stated that it is important to distinguish between the original cause of death on the death certificate and the coded cause in the vital records database as these may differ, and highlighted that vital statistics are compiled from the coded records. The authors' also noted that review procedures were variable in terms of the number of reviewers; whether the reviewers were blinded to the original death certificate; whether reviews by different reviewers were conducted independently; and if so, how disagreements between reviewers were resolved.

3.3.2 Studies identified

Our literature search for validation studies[12] focussed on general all cause mortality and some broad COD categories based on the researchers' personal knowledge of types of condition where cause of death certification is especially challenging or problematic; and/or where concerns had been raised with respect to cause of death reporting in South Africa. Using various standard data sources, we identified 16 papers on 12 studies published over a 10 year period in the following categories:

- All causes (**10** papers; **6** studies)[13 -21]
- Ill-defined (**2** studies)[22,23]
- Cardiovascular disease (CVD) (**4** studies).[24-27]

Characteristics of studies identified

The characteristics of the studies are summarised in Table 3.1. Although most studies were conducted in developed countries, several were undertaken in developing countries. Our findings were similar to Johansson *et al.*[11] as we found that there was no standardised methodology and a variety of approaches were used for the validation process. These ranged from review of medical records by one or more doctors to verbal autopsy questionnaire reviewed by a doctor. Some studies involved the completion of a cause of death certificate. Some were hospital based while others were population based.

Table 3.1 Characteristics of the studies on validation of cause of death

	General **n = 6**	**Ill-defined** **n = 2**	**Cardiovascular** **n = 4**
Countries	**1** China **1** England **1** Sweden **2** Thailand **1** USA	**1** India **1** Iran	**1** Bahrain **3** USA
Time span (based on years of death studied)	1992 – 2007	1995 – 2006	1981 – 2006
Size: Median (range)	6,280 (440 – 69,818)	1,426 & 48,357	770 (54 – 6,316)

	General **n = 6**	**Ill-defined** **n = 2**	**Cardiovascular** **n = 4**
Settings	**2** population **1** population, restricted to those who had been hospitalised in the last year of life **1** hospital **2** autopsy (both based in teaching hospitals)	**1** population **1** hospital	**1** population **1** hospital **1** hospital & population samples **1** autopsy
Original data source	**4** DNF/death certificate **2** coded COD from death certificate	**1** death certificate **1** coded COD from death certificate	**1** death certificate **3** coded COD from death certificate
Validation data source	**1** hospital records **2** hospital records & VA **1** hospital discharge diagnosis **2** autopsy report	**1** verbal autopsy **1** hospital records	**1** hospital records **3** multiple sources including medical records, autopsy reports, interviews with relatives, nursing home records, coroner's reports, interviews with doctors, medical records etc.
Validation process	**1** reviewed by a single GP **1** sequential review by 2 physicians **1** independent review by a panel of >2 physicians & consensus **1** multi-step process involving trained regional staff and a final check by a central expert validation panel **2** validation process not described in detail	**1** independent review by 2 physicians & consensus **1** hospital records reviewed by trained physicians who completed new death certificate. ACME used to determine underlying COD on original and new death certificate.	**1** independent review by 2 physicians & consensus **3** review process not described in detail

Specific findings of studies identified

A consistent finding in these studies was that the attributed underlying cause of death was frequently incorrect. However, it is not possible to draw general findings about the types of cause of death most commonly substituted for the true underlying cause of death, as many studies did not report sufficient detail and some were not large enough to do so. The autopsy studies revealed that underlying causes of death may be missed, if based on history and medical records alone. Since the autopsies were generally carried out because of uncertainties regarding the cause of death, this finding is however not generalisable. In some studies, particularly the cardiovascular disease validation studies, mis-diagnosis assigned a less

specific cause of death within the same category (e.g. ICD-9 411-414 instead of ICD-9 410), thus having limited impact on aggregated statistics. Cancers were consistently the most accurately reported causes of death. It also seemed that more accurate reporting of underlying cause of death was noted in the young adults and misreporting increased rapidly with increasing age at death.

Studies also showed that it was possible to reduce the proportion of deaths attributed to non-specific and ill-defined conditions by means of validation procedures, even in developing countries such as India and Thailand where a high proportion of deaths are attributed to non-specific and ill-defined conditions.

A set of papers on the Thailand study conducted in 2005, published in *Population Health Metrics* in May 2010, provided a broad framework for planning a large national (or multi-district, multi-province) validation study. The study was action-oriented health systems research with a capacity-building component, and the authors emphasised epidemiological implications and public health relevance in the reporting of results. The study attempted to validate a nationally representative sample of all registered deaths and the verbal autopsy instrument used. Important methodological considerations discussed in the set of papers, that were either not addressed or not described in detail in most of the other papers reviewed, include: describing sample size calculations; providing details on quality assurance procedures used in the study; discussing potential sources of bias (e.g. sampling bias, non-response bias and refusal bias) and the likely impact of these biases on the study results; providing cross-tabulations of misclassification of the underlying cause of death by ICD code (3-digit codes); and providing adjusted estimates to minimise the impact of biases. The validation of the verbal autopsy instrument highlighted errors made through the verbal autopsy, and was used to calibrate the adjustment that they made. Other important considerations discussed in the set of papers, are the impact of language and culture on the cause of death reporting process.

3.3 Lessons learnt from the literature review

In planning a validation study, it is important to specify the review process in detail and to use standard procedures. It is generally useful to have the cause of death validated independently by more than one reviewer with discrepancies resolved by consensus, a third reviewer, or a

review panel. It is also useful to blind the reviewer(s) to the original cause of death in order to minimise ascertainment bias.

Sample size and choice of study population(s) should be determined by the objectives of the study, and should be considered in the planning of validation studies. If the sample size is too small (underpowered), this limits the inferences that can be drawn from the study and limits the possibility of subgroup analyses (e.g. by specific cause and age). Sample sizes that are larger than necessary (as determined by the study objectives) waste time, money and human resources.

If the aim of the study is to improve the quality of the vital registration system, then it is necessary to address both the quality of death certificate completion and the quality of coding. Good quality of death certification without good quality coding of the underlying cause of death, or good quality coding without good quality death certification, will not translate into a good quality vital (death) registration system.

The national study undertaken by Thailand may have long term benefits in improving the system of collecting cause of death statistics, beyond providing estimates of the extent of errors. Improving the system of collecting cause of death statistics should be one of the main objectives of any validation study.

4 Conclusions and recommendation

Analysis of the 2007 StatsSA deaths data indicates that the quality of cause of death statistics has improved since the evaluation undertaken by the WHO. Less than 20% of the deaths were due to ill-defined causes of death; in 2007, 13.8% deaths were due to ill-defined natural causes and 5% were due to injuries of undetermined cause, giving an overall proportion of 18.8%. Together with improved completeness of death registration, this shifts South Africa into the category of countries with medium quality rather than poor. However, in terms of monitoring the health status of the nation and understanding the burden of disease, the extent of ill-defined causes together with 9.0% of deaths being due to garbage codes, highlights the urgent need to improve the quality of cause of death certification.

The analysis of the data has revealed several challenges that need to be addressed:

- There are multiple indications that doctors need training in cause of death certification. Furthermore, government's new response to HIV/AIDS provides an important opportunity to encourage doctors to identify HIV as the underlying cause when a patient dies from AIDS.
- There is a need to strengthen coding and application of ACME in the national statistical office.
- A large proportion of ill defined deaths were reported outside a health facility. Furthermore, $\pm$10% of deaths maybe certified by headman. It may be useful to consider a system of verbal autopsy to identify these cases. Few validation studies have been conducted. It would be useful for validation studies to be conducted in a wider range of settings and provinces.
- Currently the DNF does not make provision for the manner of death from injuries e.g. homicide, suicide or accident. The manner of death from injuries provides vital information for planning hence the DNF needs to be revised to include reporting nature of injury. This would mean that the DNF will only be regarded as completed after the inquest to determine cause of death was done and this information was added to the form.
- The 2nd National Burden of Disease study currently underway in SA will need to use methods to correct for the data inadequacies in the vita registration data so it can better inform policy making.

References

1. Mathers CD, Ma Fat D, Inoue M, Roa C, Lopez AD. Counting the dead and what they died from: an assessment of the global status of cause of death data. *Bull World Health Organ* 2005; 83(3): 171-7.
2. Groenewald P, Nannan N, Bourne D, Laubscher R, Bradshaw D. Identifying deaths from AIDS in South Africa. *AIDS* 2005; 19(2): 193-201.
3. Groenewald P, Bradshaw D, Dorrington RE, Bourne D, Laubscher R, Nannan N. Identifying deaths from AIDS in South Africa: an update. *AIDS* 2005; 19(3): 193-201.
4. Bradshaw D, Groenewald P, Laubscher R, Nannan N, Nojilana B, Norman R, Pieterse D, Schneider M. *Initial estimates from the South African National Burden of Disease Study, 2000.* MRC Report. March 2003. ISBN: 1-919809-64-3.
5. Statistics South Africa. Mortality and causes of death in South Africa, 2007: findings from death notification. Statistical Release PO309.3. Pretoria: Statistics South Africa, 2009. Available from: http://www.statssa.gov.za/publications/P03093/P030932007.pdf.
6. Burger EH, Van der Merwe L, Volmink J. Errors in the completion of the Death Notification Form. *S Afr Med J* 2007; 97: 1077-1081.
7. Nojilana B, Groenewald P, Bradshaw D, Reagon G. Quality of cause of death certification at an academic hospital in Cape Town, South Africa. *S Afr Med J* 2009; 99 (9). 648-652.
8. Burger EH, Groenewald P, Bradshaw D, Ward A, Yudkin P, Volmink J. Validation of cause of death statistics in Cape Town, South Africa. Submitted
9. Republic of South Africa. South African Government Gazette 1999. B-1680 No 20350
10. Naghavi N, Makela S, Foreman K, O'Brien J, Pourmalek F, Lozano R. Algorithms for enhancing public health utility of national causes-of-death data. *Population Health Metrics* 2010, 8:9.
11. Johansson LA, Westerling R, Rosenberg HM. Methodology of studies evaluating death certificate accuracy were flawed. *J Clin Epidemiol* 2006; 59: 125-131.
12. Metcalf C. Literature review: Cause of death validation studies. Report prepared for BOD Research Unit. 2010.
13. Choprapawon C, Porapakkham Y, Sablon O, Panjajaru R, Jhantharatat B. Thailand's national death registration reform: verifying the causes of death between July 1997 and December 1999. *Asia-Pac J Public Health* 2005; 17(2): 110-6.
14. Johansson LA, Westerling R. Comparing Swedish hospital discharge records with death certificates: implications for mortality statistics. *Int J Epidemiol* 2000; 29(3):495-502.
15. Johansson LA, Westerling R. Comparing hospital discharge records with death certificates: can the differences be explained? *J Epidemiol Community Health* 2002; 56(4):301-8.

16. Rao C, Yang G, Hu J, Ma J, Xia W, Lopez AD. Validation of cause-of-death statistics in urban China. *Int J Epidemiol* 2007; 36(3): 642-51.

17. Sington JD, Cottrell BJ. Analysis of the sensitivity of death certificates in 440 hospital deaths: a comparison with necropsy findings. *J Clin Pathol* 2002; 55(7): 499-502.

18. Rao C, Porapakkham Y, Pattaraarchachai J, Polprasert W, Swampunyalert N, Lopez AD. Verifying causes of death in Thailand: rationale and methods for empirical investigation. *Popul Health Metr* 2010; 8:11.

19. Pattaraarchachai J, Rao C, Polprasert W, Porapakkham Y, Pao-In W, Singwerathum N, Lopez AD. Cause-specific mortality patterns among hospital deaths in Thailand: validating routine death certification. *Popul Health Metr* 2010; 8:12.

20. Polprasert W, Rao C, Adair T, Pattaraarchachai J, Porapakkham Y, Lopez AD. Cause-of-death ascertainment for deaths that occur outside hospitals in Thailand: application of verbal autopsy methods. Popul Health Metr. 2010; 8:13.

21. Porapakkham Y, Rao C, Pattaraarchachai J, Polprasert W, Vos T, Adair T, Lopez AD. Estimated causes of death in Thailand, 2005: implications for health policy. *Popul Health Metr* 2010; 8:14.

22. Gajalakshmi V, Peto R. Verbal autopsy of 80,000 adult deaths in Tamilnadu, South India. *BMC Public Health* 2004; 4: 47.

23. Khosravi A, Rao C, Naghavi M, Taylor R, Jafari N, Lopez AD. Impact of misclassification on measures of cardiovascular disease mortality in the Islamic Republic of Iran: A cross-sectional study. *Bull World Health Organ* 2008; 86(9): 688-696.

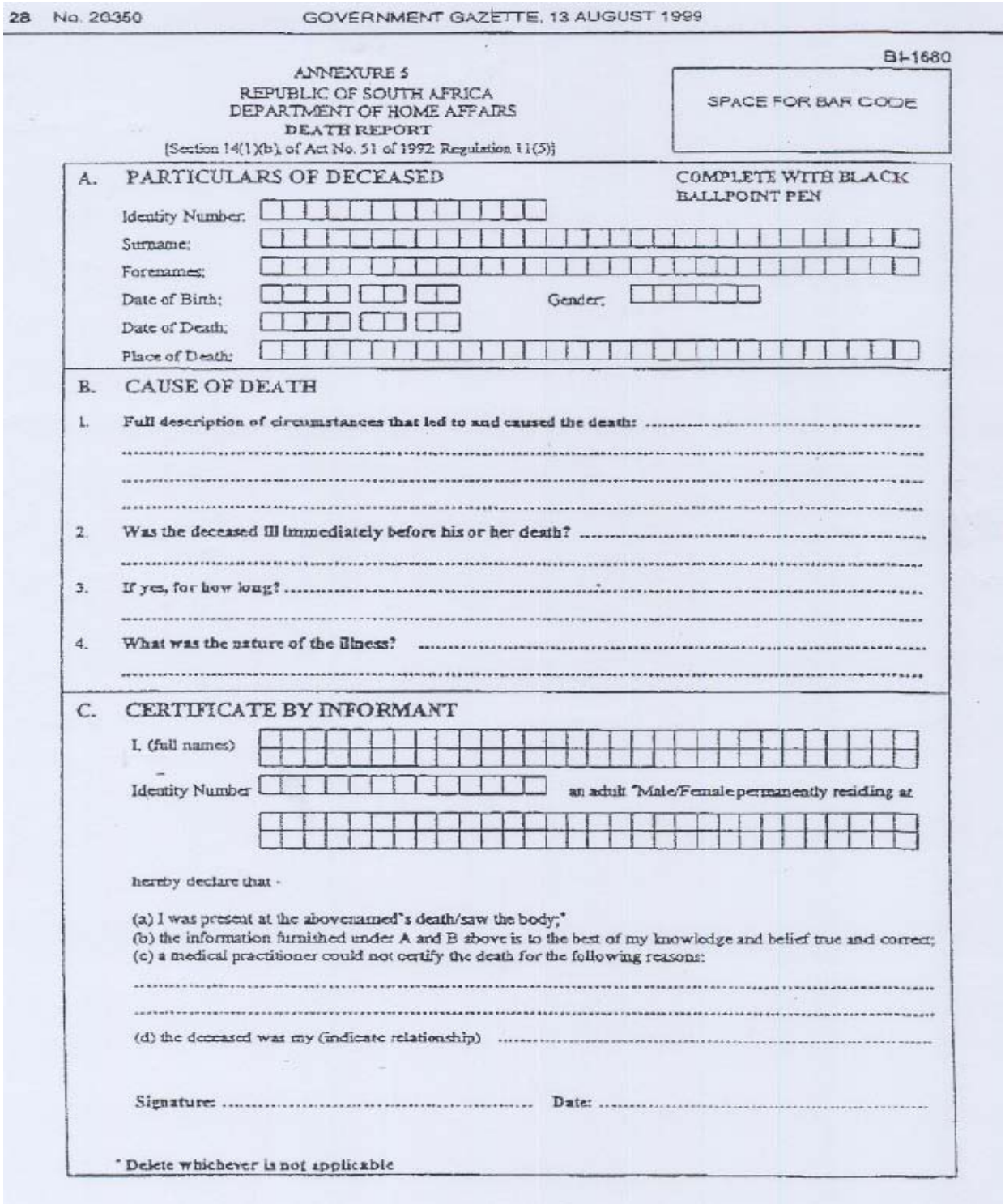
24. Al-Mahroos, R. Validity of death certificates for coding coronary heart disease as the cause of death in Bahrain. *East Mediterr Health J* 2000; 6(4): 661-9.

25. Coady SA, Sorlie PD, Cooper LS, Folsom AR, Rosamond WD, Conwill DE. Validation of death certificate diagnosis for coronary heart disease: the Atherosclerosis Risk in Communities (ARIC) Study. *J Clin Epidemiol.* 2001; 54(1): 40-50.

26. Goraya TY, Jacobsen SJ, Belau PG, Weston SA, Kottke TE, Roger VL. Validation of death certificate diagnosis of out-of-hospital coronary heart disease deaths in Olmsted County, Minnesota. *Mayo Clinic Proc*. 2000; 75(7): 681-7.

27. Tavora F, Crowder C, Kutys R, Burke A. Discrepancies in initial death certificate diagnoses in sudden unexpected out-of-hospital deaths: the role of cardiovascular autopsy. *Cardiovascular Pathol*. 2008; 17(3): 178-82.

Appendix: Form B1-1680



28 No. 20350 GOVERNMENT GAZETTE, 13 AUGUST 1999

B1-1680

ANNEXURE 5
REPUBLIC OF SOUTH AFRICA
DEPARTMENT OF HOME AFFAIRS
DEATH REPORT
[Section 14(1)(b), of Act No. 51 of 1992: Regulation 11(5)]

SPACE FOR BAR CODE

A. PARTICULARS OF DECEASED

COMPLETE WITH BLACK BALLPOINT PEN

Identity Number:
Surname:
Forenames:
Date of Birth: Gender:
Date of Death:
Place of Death:

B. CAUSE OF DEATH

1. **Full description of circumstances that led to and caused the death:** .....
2. **Was the deceased ill immediately before his or her death?** .....
3. **If yes, for how long?** .....
4. **What was the nature of the illness?** .....

C. CERTIFICATE BY INFORMANT

I, (full names)
Identity Number an adult *Male/Female permanently residing at

hereby declare that -

(a) I was present at the abovenamed*'s death/saw the body;*
(b) the information furnished under A and B above is to the best of my knowledge and belief true and correct;
(c) a medical practitioner could not certify the death for the following reasons:

(d) the deceased was my (indicate relationship) .....

Signature: ..... Date: .....

* Delete whichever is not applicable