TECHNICAL REPORT

SAMRC BURDEN OF DISEASE RESEARCH UNIT

Second National Burden of Disease Study for South Africa

Data cleaning, validation and SA NBD list

SOUTH AFRICAN MEDICAL RESEARCH COUNCIL

Second South African National Burden of Disease Study: Data cleaning, validation and SA NBD List

Victoria Pillay-Van Wyk, Ria Laubscher, William Msemburi, Rob E. Dorrington, Pam Groenewald, Theo Vos, Richard Matzopoulos, Megan Prinsloo, Beatrice Nojilana, Nadine Nannan, Nontuthuzelo Somdyala, Nomfuneko Sithole, Ian Neethling, Edward Nicol, Anastasia Rossouw, Jané D. Joubert and Debbie Bradshaw

Technical Report

Burden of Disease

Research Unit

December 2014

Copyright

ISBN: 978-1-920618-17-9

Suggested citation

Pillay-Van Wyk V, Laubscher R, Msemburi W, Dorrington RE, Groenewald P, Vos T, *et al.* *Second South African National Burden of Disease Study: Data cleaning, validation and SA NBD List*. Cape Town: Burden of Disease Research Unit, South African Medical Research Council, 2014. ISBN: 978-1-920618-17-9.
Available at: http://www.mrc.ac.za/bod/reports.htm

Acknowledgements

The South African Burden of Disease team consists of the following members (in alphabetical order) Debbie Bradshaw, Rob Dorrington, Pam Groenewald, Jané D. Joubert, Ria Laubscher, Richard Matzopoulos, William Msemburi, Nadine Nannan, Ian Neethling, Edward Nicol, Beatrice Nojilana, Victoria Pillay-Van Wyk (Project Coordinator), Megan Prinsloo, Anastasia Rossouw, Nomfuneko Sithole, Nontuthuzelo Somdyala and Theo Vos.

We thank Statistics South Africa for providing cause-of-death data and Ms Elize De Kock and Ms Claudette Garnie for administrative support on the SA NBD 2 study, and Ms Michelle Galloway and Dr Alpa Somaiya for editing the report.

Table of Contents

List of Tables

List of Figures

Acronyms

${}_4q_1$	Conditional probability of a 1-year-old dying before reaching age 5
${}_5q_0$	Probability of a newly born child dying before reaching age 5
q_0	Probability of a newly born child dying before reaching age 1
AIC	Akaike Information Criterion
ANACOD	Analysing mortality level and cause-of-death data
ASSA	Actuarial Society of South Africa
CS	Community survey
CMR	Child mortality rate
GBD	Global burden of disease
GIT	Gastrointestinal tract
GT	Gauteng
ICD-10	International statistical classification of diseases and health-related problems – 10th revision
IMR	Infant mortality rate
IMS	Injury Mortality Survey
KZN	KwaZulu-Natal
NGO	Non-governmental organisation
NBD	National burden of disease
NIMSS	National Injury Mortality Surveillance System
P-code	Perinatal code
SA	South Africa
SA NBD 2	Second National Burden of Disease Study
Stats SA	Statistics South Africa
TB	Tuberculosis
U5MR	Under-5 mortality rate

VR	Vital registration
WC	Western Cape
WHO	World Health Organization
YLDs	Years lived with disability
YLLs	Years of life lost

Glossary

ANACoD (Analysing mortality level and cause-of-death data)

This tool is intended to help build analytical capacity to assess the quality of mortality statistics to enhance their value in informing health policies and programmes. It provides relatively simple ways of analysing the internal validity and coherence of mortality data and shows how comparisons with other data sources can be used to assess consistency and plausibility. By carrying out these simple checks, data collectors and practitioners will be able to diagnose weaknesses in their data.

Broad cause groups

Type 1 (pre-transitional causes): Communicable diseases, maternal causes, perinatal conditions and nutritional deficiencies. HIV/AIDS is usually in Group I, however, in the SA NBD 2 Study, since there has been a strong shift towards an integrated approach to the challenge of HIV/AIDS and TB, HIV/AIDS and TB was treated as a separate group.

Type 2 (non-communicable diseases): These include malignant neoplasms; cardiovascular diseases; chronic respiratory diseases; digestive, musculoskeletal and genitourinary conditions; as well as mental disorders and neurological conditions.

Type 3 (Injuries): Unintentional and intentional injuries.

Cause of death

Medical certification of death involves the doctor deciding on the condition or event that resulted in the death and indicating the sequence of events in a structured, two-part format on the death-notification form. Causes are identified according to their contribution to the sequence. From a public-health perspective aimed at preventing premature mortality, the underlying cause of death is the most important.

Underlying cause of death

The disease or injury that initiated the train of morbid events leading directly to death or the circumstances of the accident or violence that produced the fatal injury. This is also called the primary medical cause of death.

Immediate cause of death

Any disease or condition entered on line (a) in Part 1 of the death certificate directly leading to death and a consequence of diseases entered on lower lines of Part 1. This is also known as the terminal, direct or final cause of death.

Intermediate cause

Any cause between the underlying cause and the immediate cause of death, or, if the certificate has not been filled out correctly, any condition that the certifier should have reported. This is also known as a complication of the underlying cause.

Contributory cause

Any cause of death that is neither the immediate, intervening, originating antecedent nor underlying is a contributory cause of death (i.e. conditions that should be reported in Part 2).

Mechanism of death

The physiological disturbance in the body at the time of death, e.g. metabolic acidosis, hypokalaemia or acute cardiac failure.

Multiple causes of death

All the diseases, morbid conditions or injuries that either resulted in or contributed to death, and the circumstances of the accident or violence that produced any such injuries.

Child mortality age groups

The age groups considered in this report are neonatal (0–27 days), early neonatal (0-6 days) and late neonatal (7–27 days), post-neonatal (1–11 months) and children (1–4 years). In terms of health programmes, a key distinction would be between newborns (<1 month) in the neonatal period, and babies and children aged 1–59 months. Demographically, the key age groups would be infants (0–11 months) measured by the infant-mortality rate (IMR), which is closely approximate to q_0, and children (1–4 years) measured by the child-mortality rate (CMR), ${}_4q_1$. These two age groups combined, give the under-5 mortality rate (U5MR), ${}_5q_0$.

Disease categories in the second SA NBD list

Cardiovascular diseases

Cardiovascular disease (CVD) covers a wide array of disorders, including diseases of the cardiac muscle and the vascular system supplying the heart, brain and other vital organs. The most common manifestations of CVD are ischaemic heart disease, congestive heart failure and stroke.

Conditions originating during the perinatal period

Deaths from these causes (primarily low birth weight and birth trauma/asphyxia) may occur at any age, but are largely confined to the perinatal period. This group of causes is distinct from perinatal mortality defined on the basis of age and includes stillbirths from 27 weeks of gestation and live births that die up to 6 days of age. Stillbirths are excluded from the SA NBD 2 list and are not part of the category of conditions originating during the perinatal period.

Congenital abnormalities

Malformations, deformities and chromosomal abnormalities that are present at birth.

Diabetes mellitus

Diabetes mellitus, or simply diabetes, is a group of metabolic diseases in which a person has high blood sugar, either because the pancreas does not produce enough insulin, or because cells do not respond to the insulin that is produced. This includes type-1 diabetes mellitus ('insulin-dependent diabetes mellitus' or 'juvenile diabetes') and type-2 diabetes mellitus ('non-insulin-dependent diabetes mellitus' or 'adult-onset diabetes').

Digestive diseases

Diseases of the digestive system such as gastric ulcers appendicitis and liver disease, excluding cancers.

Endocrine, nutritional, and other blood and immune disorders

Blood disorders such as haemolytic anaemia, disorders of the immune mechanism (non-HIV related), thyroid disorders, glucose and pancreatic regulatory disorders, endocrine gland disorders, obesity and other metabolic disorders.

Genito-urinary diseases

Renal and other urinary diseases, and gynaecological and prostatic conditions, excluding cancers.

HIV/AIDS and TB

HIV and AIDS, as well as all forms of tuberculosis, are combined as a category due to the need for integrated care for these conditions.

Infectious/parasitic diseases

Caused by the transmission of a parasite, virus, fungus or bacteria. These include sexually transmitted infections, diarrhoeal disease, vaccine-preventable childhood diseases, meningitis and encephalitis, malaria, and a few others.

Intentional injuries

Due to violence and subdivided into self-inflicted injuries, interpersonal violence, legal intervention and war-related injuries.

Malignant neoplasms

Cancer is a broad group of diseases that involve unregulated cell growth. In cancer, cells divide and grow uncontrollably, forming malignant tumours (neoplasms), and invade nearby parts of the body. The cancer may also spread to more distant parts of the body through the lymphatic system or bloodstream.

Maternal conditions

Deaths of women while pregnant or within 42 days of termination of pregnancy from any cause related to or aggravated by the pregnancy or its management, but not from accidental or incidental causes. Maternal deaths can be considered direct or indirect maternal causes.

- ***Direct maternal death:*** Deaths resulting from obstetric complications of the pregnancy state (pregnancy, labour and puerperium), from interventions, omissions, incorrect treatment or a chain of events resulting from any of the above.

- ***Indirect maternal death:*** Deaths resulting from previous existing disease, or disease that developed during pregnancy, which were not due to direct obstetric causes but were aggravated by the physiological effects of pregnancy.

Mental disorders

Mental and behavioural disorders including depression, anxiety, schizophrenia, alcohol and drug abuse, eating disorders, childhood-behaviour disorders, development disorders and mental retardation.

Musculo-skeletal disorders

Rheumatoid and osteoarthritis, back and neck pain and other musculo-skeletal disorders.

Nervous system disorders

Neuro-psychiatric conditions including Alzheimer's and Parkinson's disease, and other neurological degenerative diseases including epilepsy and other neurological conditions.

Nutritional conditions

Malnutrition resulting from energy or nutrient deficiencies, or other nutritional disorders.

Oral conditions

Dental caries, periodontal diseases and other oral conditions.

Respiratory diseases

Chronic respiratory conditions such as asthma or chronic obstructive pulmonary disease, excluding infections.

Respiratory infections

These include respiratory tract infections and otitis media (middle-ear infection). Respiratory tract infection refers to a number of infectious diseases involving the respiratory tract. These infections are further classified as upper respiratory tract infections (URI or URTI) or lower respiratory tract infections (LRI or LRTI). An example of a lower respiratory infection is pneumonia and that of an upper respiratory infection is the common cold.

Sense organ diseases

Conditions affecting sight and hearing.

Skin diseases

Skin infections and disorders, excluding cancer.

Unintentional injuries

Due to fate, accidents, or other unpredictable and uncontrollable events, and subdivided into road-traffic injuries, poisoning, falls, fires, drowning, surgical and medical misadventure, suffocation and foreign bodies, and other unintentional injuries.

Garbage codes

Global burden of disease (GBD) experts coined this term for ICD codes for mortality data that do not signify an underlying cause of death. These include ill-defined signs and symptoms, intermediate causes of death (e.g. septicaemia), mechanisms of death (e.g. cardiac arrest) or partially specified causes (e.g. cancer with unknown sites).

National burden of disease study (NBD)

A comprehensive demographic and epidemiological framework to estimate health gaps of a country for an extensive set of disease and injury causes, and for major risk factors, using all available mortality and health data, and methods to ensure internal consistency and comparability of estimates.

Years of life lost (YLLs)

The years of life lost, compared to a normative standard, due to premature mortality. This is a mortality measure of the health gap.

1 Introduction

1.1 Background

Over the past decade, the number of deaths in South Africa has increased at an alarming rate and, facing a quadruple burden of disease, the health system is under severe pressure to meet the care demands of ill health that lead to these deaths. Updating knowledge and an on-going understanding of the country's burden of disease and injury are fundamental to health planning and decision making. Such information should enable national and provincial government to begin to address and reduce the disease burden, and allocate resources more appropriately in resource-constrained settings instead of reacting to the pressures placed upon the health system. Information about mortality trends provides the first step in comprehensively understanding the burden of disease, which ideally includes the extent of the loss of health resulting from the non-fatal effects of illness and injury. Furthermore, reliable and comparable assessments of the impact of modifiable health risks, as well as information about the cost-effectiveness of interventions, are also needed to identify relevant prevention and health-promotion activities.

In South Africa, the first national burden of disease study conducted in 2000 was an important resource for government, non-governmental organisations (NGOs) and the public for guiding health and resource priorities (Bradshaw, *et al*., 2003, Bradshaw *et al*, 2005, Norman *et al*, 2007). Adopting the burden of disease approach developed by Murray and Lopez (1996), it has earned wide appreciation among local health researchers and has been used extensively to inform health-sector reviews (e.g. Chopra *et al*, 2009). The second National Burden of Disease Study (SA NBD 2) was performed to assist national and provincial governments to proactively address and reduce ill health and its risks, and to assist in measuring progress towards selected Millennium Development Goals (MDGs).

1.2 Purpose of this report

This technical report details the revised SA NBD list; the data cleaning and validation steps applied to the cause-of-death data from Statistics South Africa (Stats SA) (Stats SA, 2013a); the modelling applied to the National Injury Mortality Survey data; and, the adjustments made to the number of deaths. These analyses form the foundation of mortality estimates for the SA NBD 2.

2 Revised SA NBD list

The structure and composition of the list used in a burden of disease study has implications for the resultant rankings of diseases. For example, separating the cancers into different sites may lead to them not featuring in the top causes of death when single causes are ranked. The 2010 Global Burden of Disease (GBD) Study (Naghavi *et al*, 2010, Murray *et al*, 2012), the 2005 GBD study (GBD, 2005), the 2000 SA NBD (Bradshaw *et al*, 2003) and classifications, together with consultation with experts, were used to create the list for study (SA NBD 2). The level of disaggregation of conditions has been guided by the criteria used in the GBD lists. These include the number of deaths due to specific causes: causes that are important for public-health policy or health services and health budget, and causes that are not so broad as to be uninformative for public-health policy yet not too specific to make comparisons with other settings difficult.

The revised SA NBD list has been designed to reflect local cause-of-death patterns that differ from the GBD list. For example, indirect maternal conditions, which have increased as a result of the high proportions of women infected by HIV, have been separated from other maternal conditions. Another example is the inclusion of septicaemia in the list even though it is not an underlying cause of death. The high number of deaths reported from this cause in South Africa, together with the lack of information about the causes of septicaemia, made this a more practical approach. The perinatal conditions for the second SA NBD list differ from the first, due to an effort to follow the groupings used by Black *et al* (2010) and Lawn *et al* (2010) in their analyses of global neonatal mortality. It is anticipated that these will be more instructive for programme interventions.

The GBD studies divide the conditions into three broad cause groups:

Type 1: Communicable diseases, maternal causes, perinatal conditions and nutritional deficiencies

Type 2: Non-communicable diseases

Type 3: Injuries

Due to the size of the HIV/AIDS burden in South Africa, the initial SA NBD study introduced a fourth group by dividing Type I into two sub-groups to allow for HIV/AIDS alone. Since there has been a strong shift towards an integrated approach to the challenge of HIV/AIDS and TB, it has been decided to separate Type 1 into two sub-groups namely (a) Type 1a

maternal, perinatal, nutritional and other infectious diseases, and (b) Type 1b HIV/AIDS and TB.

Table 1 shows the national burden of disease list for the SA NBD 2 study and the respective three-digit code from the International Classification of Diseases (ICD-10) (WHO, 1992). It has evolved from the list used in the initial study, after reviewing the 2010 GBD list. As the 2010 GBD is representative of the 'global' disease burden and needs to accommodate the varied global health challenges, it has more conditions than considered appropriate for the SA NBD 2 list. The initial SA NBD list had 24 categories of diseases while the SA NBD 2 has 23 categories. Due to concern about limited access to full forensic investigation, 'Cot death' is regarded as an ill-defined condition and has been removed from the list. The 23 categories are identified by the capital letters A–W in the SA NBD 2 list. The SA NBD 2 list has 140 single causes, identified with codes ZA 1–140. There are two causes that are broken down further. ZA 132 has been divided to allow for 'Mining accidents' (ZA 132.1) to be distinguished from 'Other threats to breathing' (ZA 132.2). In addition, ZA 139 has been divided to separate 'Interpersonal violence with firearms' (ZA 139.1) from 'Interpersonal violence without firearms' (ZA 139.2).

Table 1: National burden of disease list for the SA NBD 2 study

Category	Code	SA NBD cause	ICD-10 codes
Type I: Communicable, maternal, perinatal and nutritional conditions			
A	**HIV/AIDS and TB**		
	ZA 1	Tuberculosis	A15-A19, B90, U51, U52, J90, J94
	ZA 2	HIV/AIDS	B20-B24, B33, B45, B59, C46, D84
B	**Infectious and parasitic diseases**		
	ZA 3	Sexually transmitted infections excluding HIV	A50>=1 month, A51-A53, A54>=1 month, A55-A63, N70-N73, A64
	ZA 4	Diarrhoeal diseases	A00-A09, K52
	ZA 5	Vaccine-preventable childhood	A33-A35 >= 1 month, A36-A37, A80, B01, B05-B06, B26, B91
	ZA 6	Meningitis and encephalitis	A39, G00-G03, G05, A83-A87, G04, G09, G93
	ZA 7	Hepatitis	B15-B19
	ZA 8	Septicaemia	A40-A41
	ZA 9	Malaria	B50-B54
	ZA 10	Intestinal parasites, other parasitic and vector diseases	A82, A90-A91,A95, B55-B57, B65, B67-B74, B76-B82, B89
	ZA 11	Other infectious diseases	A20-A28, A30-A32, A38, A42-A49, A65-A69, A70-A74, A75-A79, A81, A88-A89, A92-A94, A96-A99, B00, B02-B04, B07-B09, B25, B27-B32, B34, B37-B44, B46-B49, B58, B60-B64, B66, B75, B83, B92, B95-B99, B94

Category	Code	SA NBD cause	ICD-10 codes
C	**Respiratory infections**		
	ZA 12	Lower respiratory infections	J09-J18, J20-J22, P23 >6 days, J86
	ZA 13	Upper respiratory infections	J00-J06
	ZA 14	Otitis media	H65-H66
D	**Maternal conditions**		
	ZA 15	Maternal haemorrhage	O20, O44-O46, O67, O72
	ZA 16	Maternal sepsis	085-086
	ZA 17	Hypertensive disorders of pregnancy	O10-O16
	ZA 18	Obstructed labour	O64-O66
	ZA 19	Abortion	O00-O08
	ZA 20	Indirect maternal causes	O98, O99
	ZA 21	Other maternal	O21-O43, O47-O48, O60-O63, O68-O71, O73-O75, O80-O84, O87-O92, O94-O96
E	**Conditions originating during the perinatal period**		
	ZA 22	Preterm birth complications	P01, P07, P22, P23 <6 days, P25-P28, P29, P52, P61, P77, P80
	ZA 23	Birth asphyxia and trauma	P02, P03, P10-P15, P20-P21, P24, P50, P90-P91
	ZA 24	Sepsis and other newborn infectious conditions	P35, P36-P39, P60, (A33-A35, A50, A54) < 1 month
	ZA 25	Other non-infectious conditions arising in the perinatal period	P00, P04, P05, P08, P51, P53-P59, P62-P74, P76, P78-P79, P81-P89, P93-P94
F	**Nutritional deficiencies**		
	ZA 26	Protein-energy malnutrition	E40-E46
	ZA 27	Vitamin-A deficiency	E50
	ZA 28	Iron-deficiency anaemia	D50
	ZA 29	Other nutritional disorders	D51-D53, E00-E02, E51-E63
Type II: Non-communicable diseases			
G	**Malignant neoplasms**		
	ZA 30	Mouth cancer	C00-C08
	ZA 31	Oropharynx, nasopharynx and other pharynx cancers	C09-C13
	ZA 32	Oesophageal cancer	C15
	ZA 33	Stomach cancer	C16
	ZA 34	Colo-rectal cancer	C18-C21
	ZA 35	Liver cancer	C22
	ZA 36	Gallbladder and biliary tract cancers	C23-C24
	ZA 37	Pancreas cancer	C25
	ZA 38	Larynx cancer	C32
	ZA 39	Trachea, bronchi and lung cancers	C33-C34
	ZA 40	Melanoma cancer	C43
	ZA 41	Other skin cancers	C44
	ZA 42	Breast cancer	C50
	ZA 43	Cervix cancer	C53

Category	Code	SA NBD cause	ICD-10 codes
	ZA 44	Corpus uteri cancer	C54-C55
	ZA 45	Ovary cancer	C56
	ZA 46	Prostate cancer	C61
	ZA 47	Testis cancer	C62
	ZA 48	Bladder cancer	C67
	ZA 49	Kidney cancer	C64-C66, C68
	ZA 50	Brain cancer	C70-C72
	ZA 51	Thyroid cancer	C73
	ZA 52	Hodgkin's lymphoma	C81
	ZA 53	Non-Hodgkin's lymphoma	C82-C85, C96
	ZA 54	Multiple myeloma	C88, C90
	ZA 55	Leukaemia	C91-C95
	ZA 56	Other malignant neoplasms	C17, C30-C31, C37-C38, C40-C41, C45, C47-C49, C51-C52, C57-C58, C60, C63, C69, C74-C75, C77-C79, C97
H	**Benign neoplasms**		
	ZA 57	Benign neoplasms	D01-D48
I	**Diabetes mellitus**		
	ZA 58	Diabetes mellitus	E10-E14
J	**Endocrine nutritional, blood and immune disorders**		
	ZA 59	Endocrine nutritional, blood and immune disorders	D55-D62, D64, D66-D83, D86-D89, E03-E07, E15-E16, E20-E32, E34, E65-E68, E70-E80, E83-E85, E88
K	**Mental disorders**		
	ZA 60	Unipolar depressive disorder	F32-F33
	ZA 61	Bipolar affective disorder	F30-F31
	ZA 62	Schizophrenia	F20-F29
	ZA 63	Alcohol dependence	F10
	ZA 64	Drug use	F11-F16, F18-F19
	ZA 65	Anxiety disorders	F40-F44
	ZA 66	Eating disorders	F50
	ZA 67	Development disorders	F80-F82, F84, F88-F89
	ZA 68	Childhood-behaviour disorders	F90-F92
	ZA 69	Mental retardation not included as sequelae elsewhere	F70-F73, F78, F79
	ZA 70	Other mental and behavioural disorders	F04-F07, F09, F17, F34-39, F45-F48, F51-F59, F60-F69, F93-F98
L	**Nervous system disorders**		
	ZA 71	Alzheimer’s and other degenerative diseases of the nervous system	F01, F03, G30-G31
	ZA 72	Parkinson’s disease	G20-G21
	ZA 73	Multiple sclerosis	G35
	ZA 74	Epilepsy	G40-G41
	ZA 75	Migraine	G43

Category	Code	SA NBD cause	ICD-10 codes
	ZA 76	Non-migraine headache	G44
	ZA 77	Other neurological conditions	G06-G08, G10-G13, G14, G22-G25, G32, G36-37, G45-G47, G50-G62,G64,G70-G72, G90-G91, G95-G98, G80>=20 years
M	**Sense organ diseases**		
	ZA 78	Glaucoma	H40
	ZA 79	Cataracts	H25-H26
	ZA 80	Hearing loss not due to other diseases or injuries	H90-H91
	ZA 81	Other vision loss	H30-H35, H49-H54
	ZA 82	Other sense organ disorders	H00-H21, H27, H43-H47, H55-H57, H60-H61, H68-H83, H92-H93
N	**Cardiovascular disease**		
	ZA 83	Rheumatic heart disease	I00-I09
	ZA 84	Hypertensive heart disease	I11
	ZA 85	Ischaemic heart disease	I20-I25
	ZA 86	Pericarditis, endocarditis and myocarditis	I30, I31-I33, I38-I40
	ZA 87	Cardiomyopathy	I42
	ZA 88	Cerebrovascular disease	I60-I69, G81
	ZA 89	Conduction disorders and other dysrhythmias	I44, I45, I47, I48
	ZA 90	Aortic aneurism	I71
	ZA 91	Peripheral vascular disorders	I73
	ZA 92	Other circulatory diseases	I27, I28, I34-I37, I72, I77, I78, I80, I82-I84, I86-I89, I95-I98
O	**Respiratory disease**		
	ZA 93	Chronic obstructive pulmonary disease	J40-J44
	ZA 94	Pneumoconiosis	J60-J65
	ZA 95	Asthma	J45-J46
	ZA 96	Other interstitial lung disease	J84
	ZA 97	Other respiratory	J30-J39, J47, J66-J68, J70, J82, J85, J91-J92, J95, J98, J99
P	**Digestive diseases**		
	ZA 98	Peptic ulcer disease	K25-K27
	ZA 99	Appendicitis	K35-K37
	ZA 100	Intestinal obstruction and strangulated hernia	K40-K46, K56
	ZA 101	Non-infective inflammatory bowel disease	K50-K51
	ZA 102	Cirrhosis and liver disease	K70, K73, K74, K72, I85
	ZA 103	Gall bladder and bile duct disease	K80-K83
	ZA 104	Pancreatitis	K85, K86
	ZA 105	Other digestive diseases	K20-K22, K28-K31, K38, K55, K57-K63, K71, K76, K77, K90-K91, I81

Category	Code	SA NBD cause	ICD-10 codes
Q	**Genito-urinary diseases**		
	ZA 106	Renal disease	N00-N08, N10-N12, N17-N19, I12, I13
	ZA 107	Benign prostatic hypertrophy	N40
	ZA 108	Other urinary and gynaecological diseases	N13-N16, N20- N23, N25-N39, N41-N51, N60-N64, N75-N96, N97-N98
R	**Skin diseases**		
	ZA 109	Skin diseases	L00-L08, L10-L14, L20-L30, L40-L98, B35-B36, B85-B88
S	**Musculo-skeletal diseases**		
	ZA 110	Rheumatoid arthritis	M05-M06
	ZA 111	Osteoarthritis	M15-M19
	ZA 112	Back and neck pain	M47-M54
	ZA 113	Other musculo-skeletal disorders	M00-M02, M08, M10-M13, M20-M46, M60-M85, M87-M99
T	**Congenital abnormalities**		
	ZA 114	Neural tube defects	Q00, Q03, Q05
	ZA 115	Cleft lip and cleft palate	Q35-Q37
	ZA 116	Congenital heart anomalies	Q20-Q28
	ZA 117	Congenital disorders of GIT	Q38-Q45
	ZA 118	Urogenital malformations	Q50-Q64
	ZA 119	Fetal alcohol syndrome	Q86
	ZA 120	Down syndrome	Q90
	ZA 121	Other chromosomal abnormalities	Q91-Q99
	ZA 122	Other congenital abnormalities	Q01, Q02, Q04, Q06-Q07, Q10-Q18, Q30-Q34, Q65-Q85, Q87-Q89, P75
U	**Oral conditions**		
	ZA 123	Dental caries	K02
	ZA 124	Periodontal disease	K05
	ZA 125	Other oral diseases	K00-K01, K03-K04, K06-K14
			Type III: Injuries
V	**Unintentional injuries**		
	ZA 126	Road injuries	V01-V04,V06, V09-V80, V82-V85, V87, V89
	ZA 127	Other transport accidents	V05, V81, V86, V88, V91, V93-V98
	ZA 128	Poisonings (including herbal)	X40-X49, Y67
	ZA 129	Falls	W00-W19
	ZA 130	Fires, heat and hot substances	X00-X19
	ZA 131	Drowning	V90, V92, W65-W70, W73, W74
	ZA 132	Hanging, strangulation and other threats to breathing	W75-W84, Y37
	132.1	Mining accidents	W77, Y37
	132.2	Other threats to breathing	W75-W76, W78-W84
	ZA 133	Mechanical forces	W24-W34, W45-W46
	ZA 134	Exposure to natural forces	X30-X39

Category	Code	SA NBD cause	ICD-10 codes
	ZA 135	Adverse effects of medical and surgical treatment	Y39-Y66, Y68-Y84, Y88
	ZA 136	Animal contact	W53-W59, X20-X27, X29
	ZA 137	Other unintentional injuries	W20-W23, W35-W44, W49-W52, W60, W64, W85-W94, W99, X28, X50-X58, Y38
W	**Intentional injuries**		
	ZA 138	Self-inflicted injuries	X60-X84
	ZA 139	Interpersonal violence	X85-X99, Y00-Y09
	139.1	Interpersonal violence with firearm	X93-X95
	139.2	Interpersonal violence without firearm	X85-X92, X96-X99, Y00-Y08
	ZA 140	Legal intervention	Y35-Y36

3 Data integrity

3.1 Data cleaning

Checks were performed on the raw cause of death data coded to the NBD analysis codes and corrections were made to ensure consistency (Table 2). The criteria were identified in consultation with clinicians, GBD experts and disease-content specialists, as well as the checks in the ANACoD tool (WHO, 2011). Causes that were age or sex specific were checked to identify anomalies. For example, all maternal conditions were checked to ensure that they were reported for females of child-bearing age. Any maternal death reported for males or where the sex was unspecified were recoded to female deaths, and any maternal death reported for females <10 years or ≥ 55 years were recoded to ill-defined deaths. For the data-cleaning and adjustment steps, the ICD-10 causes were aggregated into 215 NBD analysis codes (Appendix A1). The analysis codes formed the basis for the NBD list (codes 1-142) and identified the garbage codes (codes 143-209) (Naghavi *et al*, 2010) that would need to be combined with other codes or redistributed across several codes. NBD analysis codes 143 and 144 represent ill-defined natural deaths (R codes in ICD-10), and analysis codes 145–209 are conditions that (i) are not cause of death, (ii) are intermediate causes of death, (iii) are immediate causes of death or (iv) are conditions that are not clearly specified. NBD analysis codes 210-215 are not garbage codes but rather conditions that may need to be reported separately for specific reasons. Although the data checks were conducted at the level of NBD analysis code, all recodes were conducted at ICD-10 level to maintain data integrity at the all levels of coding.

Table 2: Data validity checks performed on the raw data from Stats SA, 1997-2010

NBD analysis code	Disease	Criteria	Data amendment
14-19	Maternal conditions	Not in male or females <10 years of age or age >=55 years	If age <10 years, make ill-defined cause. If age >=55 years, make ill-defined cause. If male or unspecified sex, make sex female.
20-23	Perinatal conditions	Not >1 month	If age unspecified then make age = 0. If age = 0 & month >1 month then change P-code to appropriate code (& change any others to ill-defined). If age >0, then change cause to ill-defined.
28-54	Any cancers (C00-D44)	Not in <28 days	Make age unspecified.
28	Mouth	Not in <20 years	If UC ="C00C08" & age <=20, then make age unspecified

NBD analysis code	Disease	Criteria	Data amendment
30	Malignant neoplasm of the oesophagus	Not in <25 years	If UC=="C15" & age <=24, make age unspecified.
31	Malignant neoplasm of the stomach	Not in <25 years	If UC=="C16" & age <=24, then make age unspecified.
32	Malignant neoplasm of colon, rectum and anus	Not in <25 years	If (UC=="C18" & UC=="C19" & UC=="C20" & UC=="C21") & age <=24, then make age unspecified.
33	Malignant neoplasm of liver and intrahepatic bile ducts	Not in <10 years	If UC=="C22" & age <=9, then make age unspecified.
35	Malignant neoplasm of pancreas	Not in <25 years	If UC=="C25" & age <=24, then make age unspecified.
37	Malignant neoplasms of the trachea, bronchus and lung	Not in <20years	If UC=="C33" & UC=="C34" & age <=19, then make age unspecified.
38-39	Melanoma and other skin cancers	Not in <20years	If UC=="C43 | UC==C44" & age<=20 then make age unspecified.
40	Malignant neoplasm of breast	Not in <15 years	If UC=="C50" & age<=14 then make age unspecified.
41	Malignant neoplasm of cervix uteri	Not in male and not in <20 years	If UC =="C53", sex==male (& unspecified or age >=20), then make sex = female. If UC=="C53" & age<=19, then make age unspecified.
42	Corpus uteri cancer	Not in male and not in <20 years	If (UC =="C54" or UC=="C55"), sex==male (& unspecified or age >=20) then make sex= female. If (UC=="C54" or UC=="C55") & age <=19, then make age unspecified.
43	Malignant neoplasm of ovary	Not in male and not in <20 years	If UC =="C56", Sex==male (or unspecified & age>=20) then make Sex= female. If UC=="C56" & age<=19 then make age unspecified.
44	Malignant neoplasm of prostate	Not in female and not in <25 years	If UC =="C61", sex==female (& unspecified or age >=25, then make sex=male. If UC=="C56" & age <=25, then make age unspecified.
45	Testis cancer	Not in female and not in <10 years	If UC=="C62" & age <=9 then make age unspecified. If UC=="C62" & sex==female & age >=10years, then make male.
46	Malignant neoplasm of bladder	Not in <10 years	If UC=="C67" & age <=9, then make age unspecified.
50-51	Lymphomas and multiple myeloma	Not in < 10 years	If UC=="C81" | UC=="C82" | UC=="C83" | UC=="C84" | UC=="C85" |UC==" C96" & age<10 years then make age unspecified.
52	Myeloma	Not in <20 years	If UC=="C88 | UC== C90" & age <=20, then make age unspecified.
53	Leukemia	Not in <1 year	If UC==C91-C95" & age<1 years then make age unspecified.
58-68	Mental disorders (F01 - F99)	Not in <5 years	If UC ==F01 to F99 & age <5years, then take to Q86.
61-62	Mental and behavioural disorders due to substance abuse	Not in <10 years	If baby <1 month then code to P04. If age >1 month & <10 years, then make ill-defined.

NBD analysis code	Disease	Criteria	Data amendment
69	Alzheimer's and other degenerative diseases of nervous system	Not in <19 years	If UC=="G30" & age <=19, then make ill-defined natural.
70	Parkinson's disease	Not in <14 years	If UC=="G20" or UC=="G21" & age <=14, then make ill-defined natural.
73, 74, 76-78	Migraine, non-migraine headache, glaucoma, cataracts, hearing loss not due to other diseases, hearing loss not due to other diseases or injuries	Not causes of death	Goes to ill-defined. If NBD code==73, 74, 76-78 & age <27 days then UC=="P96". If NBD code==73, 74, 76-78 & age >27 days then UC=="R99"
82	Hypertensive disease	-	I10 will be recoded to ill-defined if it appears with other conditions.
81-90	Diseases of the circulatory system (I00-I99)	Not in <1 month	If NBD code==86 & age <1 month, then UC=="P52" otherwise goes to ill-defined.
91	Chronic obstructive airways disease and emphysema	Not in <10 years	If UC >="J40" & UC <="J44" & age <=9, then make UC="J84" Interstitial lung disease).
92	Pneumoconiosis	Not between 27 days and 15 years	Goes to ill-defined natural.
100	Cirrhosis and liver disease	Alcoholic liver disease not in <19 years	If UC=="K70" & age <=19, then make ill-defined natural.
105	Benign prostatic hypertrophy	Not in female and not in <30 years	If UC =="N40" & sex==female or unspecified & age >=30, then make sex= male. If UC=="N40" & age <=29 then make ill-defined natural.
108	Rheumatoid arthritis	Not in <5 year	If UC== "M05 | M06" & age <5 years, then make ill-defined natural.
111	Systemic lupus erythematous (M32)	Not in <15 year	If UC=="M32" & age <15 years then make ill-defined natural.
111	Polyarthritis nodosa and related conditions (M30)	Not in <15 year	If UC=="M30" & age <15 years, then make ill-defined natural.
114-120	Congenital malformations, deformations and chromosomal abnormalities	Unspecified age==0	If age==999 & NBD code >=114 & NBD code <=120 then age = 0.
112	Neural tube defects	Not in >=60 years	If UC=="Q00" or UC=="Q03" or UC=="Q05" & age >=60, then make ill-defined.
113, 115, 117	Cleft lip and cleft palate, congenital disorders of GIT, fetal alcohol syndrome	Not in >=30 years	If UC>="Q35" & UC<="Q45" & age >=30, then make ill-defined. If UC=="Q86" & age >=30, then make ill-defined.
148	Mental disorder, not otherwise specified	Not between 27 days and 5 years	Make ill-defined.
172	Adult respiratory distress syndrome	Not in <5 years	Make ill-defined.

NBD analysis code	Disease	Criteria	Data amendment
13, 55, 72, 73, 82, 92, 94, 97, 101, 109, 110, 140, 148, 175	Otitis media, other neoplasms, epilepsy, migraine, hypertensive heart disease, pneumoconiosis, other interstitial lung disease, appendicitis, gall bladder and bile duct disease, osteoarthritis, back and neck pain, self-inflicted injuries, mental disorder, not otherwise specified, pyothorax	Not in <27 days	Make P96.

3.2 Data recodes

3.2.1 Inconsistent underlying causes of death

Diabetes

Detailed analysis of the 2007 Stats SA cause-of-death data identified some inconsistencies in selection of the underlying cause of death from the reported multiple causes, including diabetes (Bradshaw *et al*, 2011). Further analysis of multiple-cause data by Nojilana *et al* (2013) revealed that for all the deaths reported with diabetes listed in Part 1 of the death notification form between 1997 and 2007, an overall proportion of 54.8% deaths also had a cardiovascular disease reported in Part 1. There were 33.4% that had hypertensive heart disease reported in Part 1, 15.3% had stroke reported in Part 1 and 11.0% had ischaemic heart disease reported in Part 1. When diabetes is reported in Part 1, it is coded as the underlying cause of death. However, given the high proportions, it is not clear if the certifying doctors intended to identify diabetes as the underlying cause.

In consultation with local diabetes and cardiovascular experts, and the GBD experts, it was decided to recode deaths recorded with diabetes as the underlying cause based on the multiple cause information in the following way:

a) Recode underlying cause to stroke (I60-I69) if stroke (I60-I69) was also reported in Part 1.

b) Recode underlying cause to ischaemic heart disease (I20-I25) if ischaemic heart disease (I20-I25) was also reported in Part 1.

c) Recode underlying cause to hypertensive heart disease (I11) if hypertensive heart disease (I11) was also reported in Part 1.

d) Recode underlying cause to hypertensive renal disease (I12) if hypertensive renal disease (I12) was also reported in Part 1.

Epilepsy

Analysis of the 2007 data also revealed cases with epilepsy as the underlying cause with stroke reported in Part 1, implying that epilepsy was the cause of a stroke. These cases were recoded to an underlying cause of stroke as the reported sequence was not considered plausible.

Hypertension

Essential hypertension (I10) was recoded to ill-defined natural (R99) if there was another cause of death reported in Part 1 as there were generally a wide range of varying causes.

3.2.2 Incorrect P-codes

The majority of children aged between 1 and 11 months were inappropriately coded by Stats SA to the Perinatal Codes (P-codes) prior to 2006 (Nannan *et al*, 2012; Stats SA 2008). Instead of using the age cut-off of 1 week, it appears that the coders used the incorrect age cut-off as under 1 year. In order to set up comparable data to evaluate trends, the P-codes for children aged 1-11 months were recoded. This coding practice was rectified from 2007 onwards (Table 3). For example, any cases of P05 'Slow fetal growth and fetal malnutrition' in age group 1–11 months were recoded to E46 'Unspecified protein-energy malnutrition'. While it is possible for perinatal conditions to result in death at an older age, the stringent recode was done to enable comparison over time.

Table 3: Recode of P-codes for child deaths 1-11 months old, 1997-2010

P-code	Cause of death	ICD-10 code	Cause of death
P05	Slow fetal growth and fetal malnutrition	E46	Unspecified protein-energy malnutrition
P22	Respiratory distress of newborn	J18	Pneumonia, organism unspecified
P23	Congenital pneumonia	J18	Pneumonia, organism unspecified
P24	Neonatal aspiration syndromes	J98	Other respiratory disorders
P25	Interstitial emphysema and related conditions originating in the perinatal period	J98	Other respiratory disorders

P-code	Cause of death	ICD-10 code	Cause of death
P28	Other respiratory conditions originating in the perinatal period	J18	Pneumonia, organism unspecified
P29	Cardiovascular disorders originating in the perinatal period	Q24	Other congenital malformations of the heart
P35	Congenital viral diseases	B24	Unspecified human immunodeficiency virus (HIV) disease
P36	Bacterial sepsis of newborn	A41	Other septicaemia
P37	Other congenital infectious and parasitic diseases	A19	Miliary tuberculosis
P39	Other infections specific to the perinatal period	B99	Other and unspecified infectious disease
P52	Intracranial non-traumatic haemorrhage of fetus and newborn	I64	Stroke, not specified as haemorrhage or infarction
P58	Neonatal jaundice due to other excessive haemolysis	R99	Other ill-defined and unspecified causes of mortality
P59	Neonatal jaundice from other and unspecified causes	R99	Other ill-defined and unspecified causes of mortality
P60	Disseminated intravascular coagulation of fetus and newborn	R99	Other ill-defined and unspecified causes of mortality
P61	Other perinatal haematological disorders	D64	Other anaemias
P70	Transitory disorders of carbohydrate metabolism specific to fetus and newborn	E15	Non diabetic hypoglycaemic coma
P74	Other transitory neonatal electrolyte and metabolic disturbances	R99	Other ill-defined and unspecified causes of mortality
P76	Other intestinal obstruction of newborn	R99	Other ill-defined and unspecified causes of mortality
P77	Necrotising enterocolitis of fetus and newborn	K52	Other non-infective gastroenteritis and colitis
P78	Other perinatal digestive system disorders	A09	Diarrhoea and gastroenteritis of presumed infectious origin
P81	Other disturbances of temperature regulation of newborn	R99	Other ill-defined and unspecified causes of mortality
P90	Convulsions of newborn	R99	Other ill-defined and unspecified causes of mortality
P91	Other disturbances of cerebral status of newborn	R99	Other ill-defined and unspecified causes of mortality
P92	Feeding problems of newborn	R99	Other ill-defined and unspecified causes of mortality
P96	Other conditions originating in the perinatal period	R99	Other ill-defined and unspecified causes of mortality
All other P-codes		R99	Other ill-defined and unspecified causes of mortality

4 Data adjustment

4.1 Data anomalies

The review of child death data indicated unusually high numbers of deaths in the months April to June of 1998 (Nannan *et al*, 2012). Examination of the all-age deaths indicated similarly high numbers of deaths in the months April to June of 1998 that could not be explained by a coherent cause or provincial pattern (Figure 1). This data anomaly was treated as an increase in completeness rather than a change in death rates (see Section 2.3 in Appendix 2).

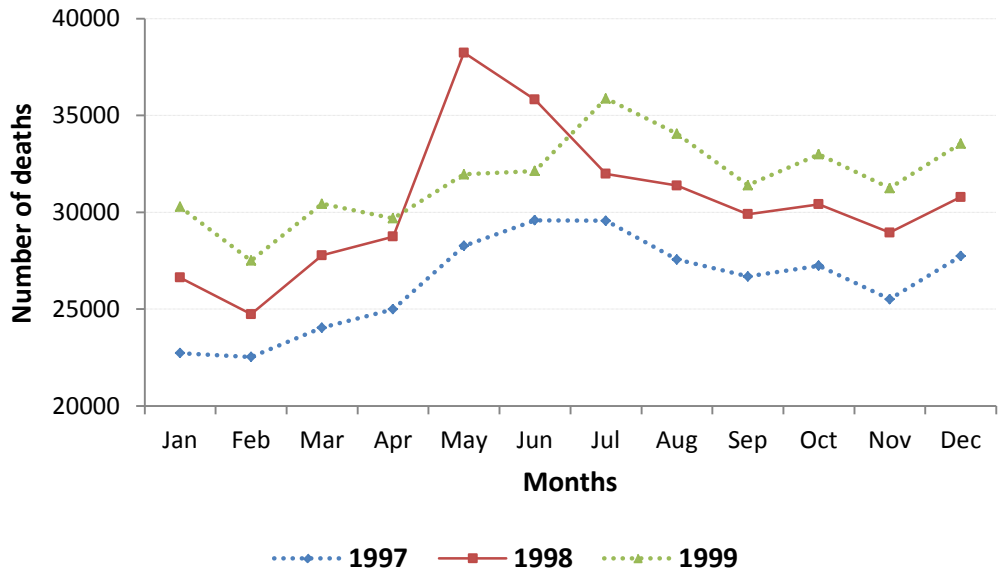

Figure 1: Unadjusted Stats SA deaths by month 1997-1998

The review of child death data also noted an unexplained increase in ill-defined natural deaths in 2004 and 2005 for children aged 1−4 years (Figure 2). In the analysis, the ill-defined deaths (NBD analysis code 143), for children aged 1−4 years, were adjusted for 2004 to follow linear change between 2003 and 2006.

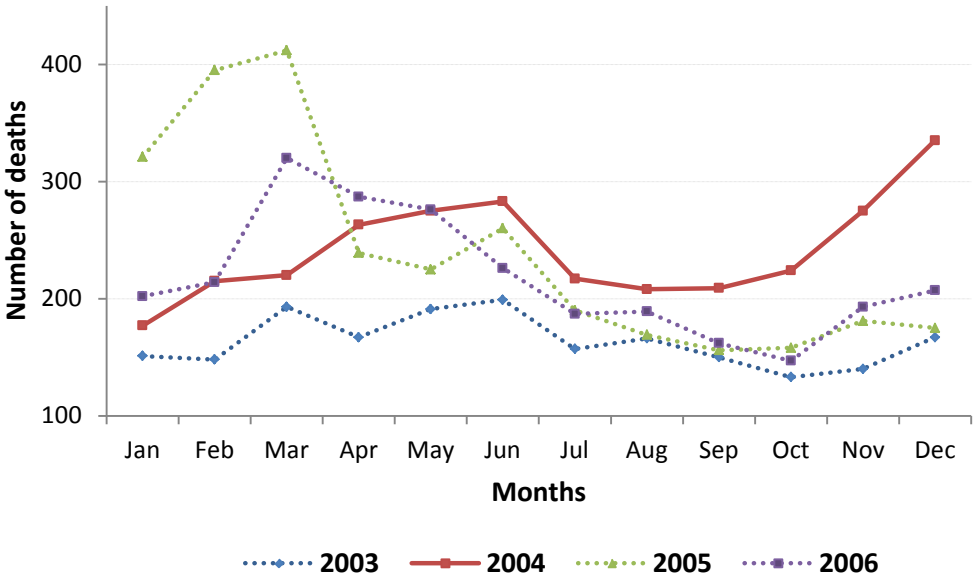

Figure 2: Unadjusted ill-defined Stats SA deaths 2003–2006

4.2 Unknown age, sex and population group

A proportional redistribution was performed of unknown sex, unknown age and unknown population group in the Stats SA database. As population group was very poorly reported prior to 1999, an average of the 1999 and 2000 proportions was applied to the 1997 and 1998 unknown population group.

4.3 Adjustment for completeness of death registration[1]

4.3.1 Estimation of completeness

A multi-stepped approach (described in detail in Appendix 2) was used to estimate the completeness of death registration for this study. The broad approach was to compare the registered deaths against empirical estimates of the mortality rates derived from surveys and censuses for infants and children under 5 years (and hence children 1-4 years), and to apply

[1] Completeness was estimated prior to the release of the 10% unit record sample of the census (Stats SA 2014a), the re-editing of the deaths reported by households (Stats SA 2014b) and the 2011 cause of death report (Stats SA 2014c). It is estimated that had the corrected data on deaths reported by households been used, infant (and hence also under-5) mortality would have increased by 3 per mille in the Western Cape and reduced by 2 per mille in the Free State and that the change would not be more than 1 per mille for the other provinces. The impact on adult mortality is likely to be negligible except for about a 1% increase in rates in Limpopo.

death distribution methods to the registered deaths of males and females aged 15 and older. Completeness for children aged 5-14 years was then estimated using linear interpolation between the completeness estimate for children aged 1-4 years and adults aged 15+ years. The resulting completeness factors for each category are shown in Table 4. The same approach was applied for each province, and the provincial completeness factors were then rescaled each year in proportion to ensure that the sum of the number of deaths adjusted for incompleteness for the provinces was equal to the total number of deaths adjusted for incompleteness for the country as a whole. Since the completeness of non-natural and natural deaths probably differs, the completeness of non-natural deaths was estimated independently (see 4.3.7) and differenced from the overall completeness to estimate the completeness of deaths from natural causes.

Table 4: National completeness of death registration (%) by age category, 1997-2010

Year	<1 year	1-4 years	5-14 years	15+ years	All ages
1997	54	40	61	80	75
1998	62	49	68	86	81
1999	62	48	67	84	80
2000	62	47	67	86	81
2001	62	46	67	87	83
2002	66	49	69	89	84
2003	71	53	71	90	86
2004	76	57	73	91	87
2005	80	59	75	91	89
2006	85	63	78	92	90
2007	85	62	78	93	90
2008	85	63	78	93	91
2009	85	62	79	93	91
2010	85	64	80	93	91

4.3.2 Infants and children under 5 years

Initial estimates of q_0 and ${}_5q_0$ were derived from several sources, including the deaths reported by households (2001 and 2011 censuses, and the 2007 Community Survey) produced by Dorrington, Moultrie and Timæus (2004), Darikwa and Dorrington (2011) and Stats SA (2013). More robust estimates of q_0 and ${}_5q_0$ were derived for 1996, 2001, 2006 and 2011 by regressing the initial provincial estimates for each year on each of the other years. Finally, the provincial estimates were scaled so that for each year, the sum of the expected number of deaths in the provinces (under the age of 1 and under the age of 5 separately) was equal to the

expected number of deaths based on the national estimates. Further details about the estimation can be found in Appendix 2.1.

The number of births by calendar year was estimated as the number required to produce the number of surviving children counted at each age at the time of the 2011 census. Further details can be found in Appendix 2.2.

The completeness of registration of deaths for individual years between the years of the point estimates of the expected number of deaths was estimated. In general,[2] this was done by assuming that completeness changed linearly with time. A few exceptions were made where either the rate of change in completeness or the mortality rates suggested that minor adjustments were necessary to make the estimates more plausible. Completeness of reporting of childhood (age 1-4) deaths was derived from the differences between reported and expected deaths under the age of 5 and under the age of 1. Further details can be found in Appendix 2.3.

4.3.3 Adults (15+ years)

Completeness of adult death registration, nationally by sex, for 1996 to 2001, was initially set to the estimates from Dorrington *et al* (2004). Initial estimates for 2001 to 2007 were derived by assuming that they changed linearly with time starting from the estimate for 2001 from Dorrington *et al* (2004) and with an average equal to that produced by applying death-distribution methods to registered deaths and the census/survey populations. After this, logistic curves (one each for each sex) were fitted to these estimates with an asymptote of 95%. As the curves for each sex crossed over, and it was difficult to make a case for separate levels of completeness of reporting deaths of men and women, it was decided to blend the curves such that completeness reached 93% by 2009.

The death-distribution method could not be applied to an intercensal period up to the 2011 census, as the registered deaths were only available up to 2010. However, a check was conducted using an estimate of the number of these deaths based on data from the rapid

[2] There were one or two years where this assumption implied implausible change in rates between one year and the next, in which case the change in the reported number of deaths was assumed to be due, in the main, to a change in completeness rather than a change in the mortality rate.

mortality surveillance system for 2011, which confirmed the estimates for the later years (see Appendix 2.4).

4.3.4 Provincial completeness of reporting adult deaths

To estimate completeness of reporting deaths provincially, it was necessary to use a different approach because the death-distribution methods would require detailed knowledge about inter-provincial migration patterns over the full period. Deaths reported by households from censuses and surveys were used to provide point estimates of the mortality rates for the age group 15+ for the provinces once they had been adjusted to account for the under-representation of deaths in older ages. Since the 1996 census did not ask households to report on deaths, the rate for 1996 was set to that from the ASSA2008 model. As the four-point estimates failed to produce convincing trends over time, they were only used to determine the overall level with the trend over time being obtained from the trend projected by the ASSA2008 model.

Initial estimates of completeness of reporting deaths were then estimated as the ratio of the rates for the age group 15+, produced by dividing unadjusted numbers of registered deaths by the mid-year population of age 15+[3] by the rates produced above. Final completeness estimates were derived by averaging the estimates for men and women, and then rebalancing them by province so that the sum of completely reported deaths by province for each year was equal to the number of completely reported deaths in this age group for the country as a whole. The method is described in further detail in Appendix 2.5.

4.3.5 Children 5-14 years of age

Completeness of reporting by single years of age for ages 5-14 were derived on the assumption that the average of completeness for ages 1-4 was equal to that estimated for the age group 1-4 years in total, and that completeness changes linearly with age between ages 1 and 15.

[3] Alternative mid-year population estimates (Dorrington, 2013).

4.3.6 Limitations

The data used to produce the above estimates suffer from a number of limitations. First, estimation of the completeness of reporting infant and child deaths requires accurate estimates of infant and under-five mortality rates at a number of time points over the period. The estimates derived from household deaths as reported in censuses and surveys is of unknown accuracy. In particular, the national estimates for 2011 appear to be too high. These problems are magnified and compounded for the provinces by the complication of porous boundaries (a birth in one province may die in another province), which applied to most provinces between 2001 and 2011. This meant that reasonableness of the magnitude of the estimates of completeness could not be used as a control on the process provincially.

Second, estimation of adult mortality would have been improved had data on the numbers of deaths for 2011 been available (in particular the registered deaths and deaths reported by households in the 2011 census by population group for each province). In addition, there is a great deal of uncertainty about the accuracy of the numbers of deaths, given that 25% were missing age data (and many missing sex data as well). Such a high number suggests the possibility of data-capturing problems. Adult deaths were also impacted by the changing boundaries.

4.3.7 Completeness for non-natural deaths

Comparison of the numbers of non-natural deaths reported in the rapid mortality surveillance from the national population register (Bradshaw *et al*, 2012), the Stats SA data and the 2009 Injury Mortality Survey (Matzopoulos *et al*, 2012) indicates under registration of non-natural deaths in both Stats SA and rapid mortality surveillance (Bradshaw *et al*, 2011) for 2009 (Figure 3). The comparison also raises concern about the number of non-natural deaths reported by Stats SA for 2009. These seem out of kilter with the deaths recorded in the population register, which the Stats SA deaths closely tracked during the preceding period. However, as the marked drop in non-natural mortality reported in the Population Register from 2008 to 2011 occurred in provinces with metropolitan areas (KwaZulu-Natal, Gauteng and Western Cape), it is likely that this trend reflects a drop in non-natural mortality.

The ratio of Stats SA numbers to IMS numbers on non-natural deaths for 2009 by age group (Figure 4) indicates that, with the exception of infants and people aged 60+ years, Stats SA recorded fewer non-natural deaths than estimated by the IMS.

In the case of infants, it is likely that the IMS under-represented non-natural deaths through inaccurate distinction between still births and deaths of live births. A national study of child homicide (Mathews et al, 2012) conducted in a sample of mortuaries estimated a similar number of non-natural infant deaths to those reported by Stats SA. The reason for the higher number of non-natural deaths in the Stats SA data for people aged 60+ years is not clear but may be a result of age-misreporting.

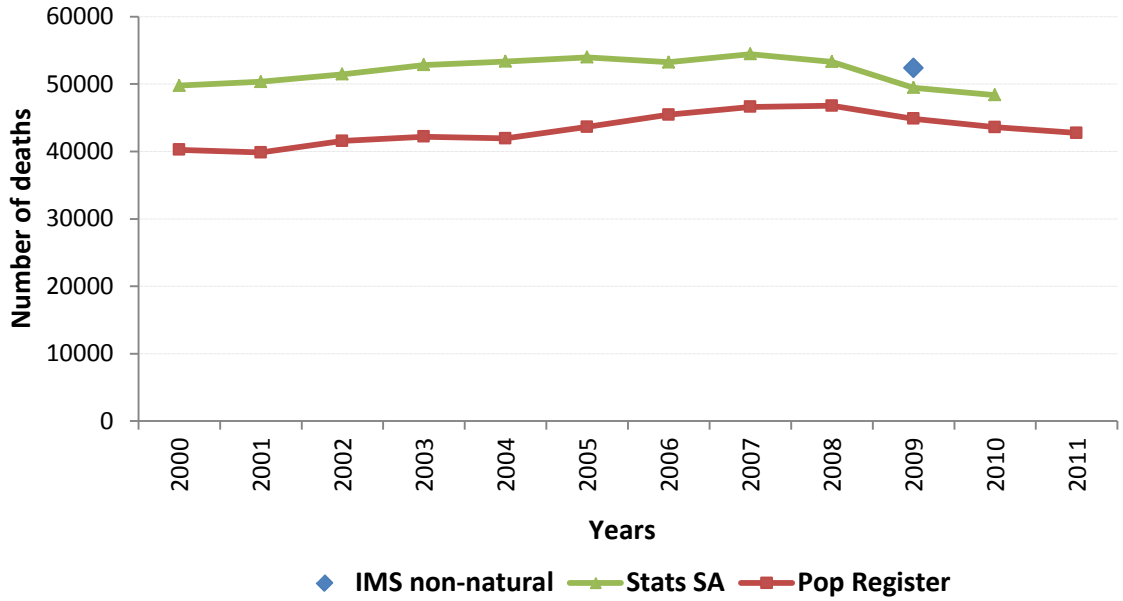

Figure 3: Number of non-natural deaths reported in the IMS, Stats SA and the Rapid Mortality Surveillance

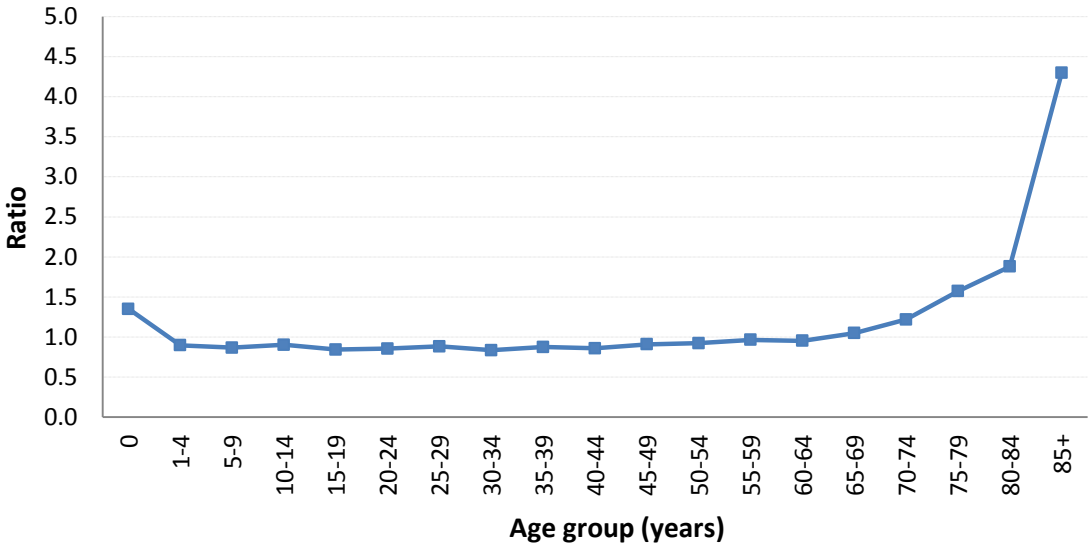

Figure 4: Ratio of Stats SA non-natural to IMS deaths by age group, SA 2009

It was decided that, with the exception of non-natural deaths in the age group <1 year, the completeness of non-natural deaths be estimated for 2009 from the national Injury Mortality Survey (IMS) of 2009. Half of the undetermined deaths observed in the IMS (*N*=868) were assumed to be and added as non-natural deaths (*N*=52 503), resulting in a total of *N*=53 371. Infant deaths were fixed to the number of non-natural deaths reported in the Stats SA data (*N*=1 043). After using a polynomial function to smooth the total non-natural deaths by age, sex and province to so as to allow for sampling errors in the data, we assumed a total of 53 629 non-natural deaths in 2009.

The completeness of non-natural deaths for other years was estimated by assuming that this followed the same rate of change as the completeness for all causes by province, age and sex. Thus the completeness of reporting of non-natural deaths was a function of the completeness estimated for 2009 and the ratio of the completeness for all causes of death in a particular year relative to the completeness for all causes in 2009.

4.4 Estimation of cause of injury (non-natural) deaths

The 2009 IMS data were used to estimate the injury-cause fractions by province, age, sex and population group for 2009. A multinomial regression with five major causes of injuries as the dependent variable, i.e. firearm homicides, non-firearm homicides, suicides, transport and other unintentional injuries, was applied to obtain smoothed estimates accounting for sampling variation. The independent variables included province, age, sex, locality type (metro vs. non-metro) and population group. The model was then used to predict the proportions of the non-natural deaths in each of the major injury cause groups by age, sex, province and population group. All cases with unknown information and undetermined cause of deaths (10.5%) were dropped from the analysis. The analysis was performed on *N*=22 399 cases. The sampling weights were used in the analysis (weighted *N*=48 322).

Box 1 shows the models with the outcome ‘causeoutc’ as a function of the independent terms, two-way and three-way interactions between sex, age in 10-year age bands, population group, province and whether the place of death was a metropolitan area, that were considered. The saturated model with a four-way interaction term was not considered in the model selection process because the purpose was to smooth the sampling variations. The interaction between age and province was not considered because the large number of categories resulted in sparse data. The Akaike Information Criterion (AIC) was used as a measure of the relative goodness

of fit of the statistical model, as it provides a tradeoff between bias and variance in model construction (Box 2). Model 7 had the lowest AIC and was therefore identified as the best-fit model.

Box 1: Model selection

Model 0: mlogit causeoutc i.Sex i.Age10 i.Race i.Province i.Metro_Non Metro

Model 1: mlogit causeoutc i. Sex *i. Age10 i.Sex*i. Race i. Sex *i. Province i. Sex *i. Metro_Non Metro

Model 2: mlogit causeoutc i. Sex *i. Race i. Sex *i. Province i. Sex *i. Age10*i. Metro_Non Metro

Model 3: mlogit causeoutc i. Sex *i. Age10 i. Sex *i. Province i. Sex *i. Race *i. Metro_Non Metro

Model 4: mlogit causeoutc i.Race*i. Age10 i. Sex *i. Province i. Sex *i. Race *i. Metro_Non Metro

Model 5: mlogit causeoutc i. Province *i.M i. Province *i. Sex i.M*i. Race i. Metro_Non Metro *i. Age10 i. Sex *i.A i. Sex *i. Race *i. Metro_Non Metro

Model 6: mlogit causeoutc i. Province *i. Metro_Non Metro i. Province *i. Sex i. Sex *i. Age10*i. Metro_Non Metro

Model 7: mlogit causeoutc i. Sex *i.A i. Sex *i. Province i. Sex *i. Metro_Non Metro i. Sex *i. Race i. Race *i. Age10 i. Race *i. Province i. Race *i. Metro_Non Metro

mlogit = logistic regression of using a categorical dependent variable; **causeoutc** = dependent outcome variable for injury cause (firearm homicide, non-fiream homicide, suicide, transport and other unintentional; **Age10** = age groups in 10-year age bands; Sex = male or female; **Race** = African, white, Indian, coloured; **Province** = nine South African provinces; **Metro_Non Metro** = the locality type of place of death being metropolitan or not metropolitan; ***** = the interaction between two variables e.g. i.Sex*Age10 is the interaction between sex and age group.

Box 2: Measure of the relative goodness of fit

Model	Obs	ll(null)	ll(model)	df	AIC
Model 0	23171	-37298.24	-34497.62	109	69213.25
Model 1	23171	-37298.24	-34350.60	208	69117.20
Model 2	23171	-37298.24	-34284.08	285	69138.16
Model 3	23171	-37298.24	-34289.42	238	69054.85
Model 4	23171	-37298.24	-34153.80	313	68933.60

Obs = number of observations; ll(null) = log likelihood under the null hypothesis (none of the coefficients are significant and the model constitutes the constant alone); **ll(model)** = log likelihood if the model is used; **df**= degrees of freedom; **AIC** = Akaike Information Criterion.

The proportion determined for the five major causes of non-natural deaths from the best-fit multinomial model, Model 7, was applied to the total number of adjusted non-natural deaths by age, sex, province and population group for the year 2009. The national breakdown on the

other unintentional deaths by age and sex was then applied to each province and population group.

Due to the lack of national data on the cause of injury deaths for earlier years, it was decided to use the NIMSS cause profile for 2000/1 as in the first NBD study and to apply linear interpolation to the proportions for each major cause by age, sex, province and population group between 2000 and 2009 to estimate the proportions for other years. The estimated proportions were normalised and applied to the observed injury data by year to estimate the number of deaths by population group, cause, province, age and sex. While the NIMSS data provide a sub-set of the injuries that occurred in 2000/1, in the absence of other information, it has been assumed that it is nationally representative of the injury-cause profile. Assuming also that the provincial and population group differentials in injury-cause profiles experienced in 2009 have remained constant, the number of injury deaths from NIMSS in 2000/1 were adjusted for each province and population group by apportioning the estimated national number of deaths from the particular cause in the same proportion as in 2009 IMS study and adjusting for any relative population change. Thus setting:

$$D_{iju2000} = \frac{D_{u2000}}{D_{u2009}} * \frac{N_{ij2000}}{N_{ij2009}} * D_{iju2009}$$

where D_{ijuy} is the number of injury deaths for province *i*, population group *j*, broad injury cause *u* and year *y*, while N_{ijy} is the population for province *i*, population group *j*, and year *y*. Injury-cause fractions could be determined for the year 2000 by province, population group, age and sex.

4.5 Redistribution of garbage and ill-defined causes

Garbage codes include conditions that are (i) not cause of death, (ii) intermediate causes of death, (iii) immediate causes of death, or (iv) conditions that are not clearly specified.

The sequential redistribution of garbage and ill-defined causes, from well-defined conditions to very vague conditions, was conducted according to Table 4. An Excel template was developed to redistribute the garbage and ill-defined codes. This was then developed in Stata 12 code to do the redistribution by age, sex, population group and province following Steps 1-6.

Step 1: For some conditions, redistribution was not required, as the garbage condition was added to an underlying cause of death.

Step 2: More unspecified garbage conditions were proportionally redistributed by age and sex across multiple conditions within the same category of disease.

Step 3: Atherosclerosis (hardening of the arteries) was redistributed proportionally by age and sex to ischaemic heart disease, stroke, aortic aneurism and peripheral vascular disease.

Step 4: Essential hypertension (persistent and pathological high blood pressure for which no specific cause can be found) was redistributed proportionally by age and sex to Hypertensive heart disease (NBD analysis code 82), Hypertensive renal and hypertensive renal and heart disease included in Renal disease (NBD analysis 104), Ischaemic heart disease (NBD analysis code 83), Cerebrovascular disease (NBD analysis code 86) and Aortic aneurysm (NBD analysis code 88).

Step 5: All ill-defined cardiovascular conditions were redistributed proportionally by age and sex to all cardiovascular. All ill-defined cancer conditions were redistributed proportionally by age and sex to all cancers.

Step 6: Heart failure was added to pulmonary oedema and redistributed proportionally by age and sex to rheumatic heart disease, hypertensive heart disease, IHD, pericarditis/endocarditis/myocarditis, cardiomyopathy, conduction disorders and other dysrhythmias. Congenital heart anomalies was added to the potential causes for <15 year olds.

Step 7: All ill-defined conditions were redistributed proportionally by age and sex to all natural deaths.

Note: Since the injury deaths have been estimated from other data sources, it was not necessary to redistribute the injury garbage codes in the Stats SA data.

Appendix A1 reports the mapping from NBD analysis codes to ICD-10 codes to NBD ZA codes, including the target ZA codes for the redistributions of garbage codes.

Table 5: NBD analysis codes and redistribution of garbage and ill-defined causes of death

NBD analysis codes	Single disease	ICD-10 codes	Target NBD analysis codes	Redistributions	Sequence for redistribution	
					Natural deaths	Injury deaths
143	Ill-defined natural: redistribute all causes	R00-R01, R03-R17, R19-R23, R25-R54, R56, R59-R63, R68, R70-R99	1-123	Proportionally redistributed by age and sex to all natural deaths. Denominator (1-123)	6	
144	Ill-defined: redistribute on some causes	R55, R57, R58, R64	143	Combine with 143 and proportionally redistributed by age and sex to all natural deaths. Denominator (1-123)	6	
145	Malignant neoplasm without specification of site	C80, C76	28-55	Proportionally redistributed by age and sex to all cancer deaths Denominator (28-55)	4	
146	Ill-defined descriptions of heart disease and unspecified disorders of circulatory system	I51, I99	81-85; 87, 90 and 86, 88, 104	Proportionally redistributed by age and sex to rheumatic heart disease, hypertensive heart disease, IHD, pericarditis, etc., cardiomyopathy, conduction disorders and other dysrhythmias and other circulatory disease. Denominator (81- 85, 87, 90)	4	
147	Hypertension only	I10	82-83, 86, 88, 104	Proportionally redistributed by age and sex to hypertensive heart disease, hypertensive renal and hypertensive renal and heart disease included in renal disease, ischaemic heart disease, cerebrovascular disease, and aortic aneurysm. Denominator (82, 83, 86, 88, 104)	4	
148	Mental disorder, not otherwise specified	F99	58-68	Proportionally redistributed by age and sex to Neuropsychiatric conditions. Denominator (58-68)	2	
149	Perinatal death unspecified cause	P96, P92, G80 <20 years	20-23	Proportionally redistributed by age and sex to pre-term birth complications, birth asphyxia and trauma, sepsis and other infections of the new born, other non-infectious conditions arising in the perinatal period. Denominator (20-23)	2	
150	Fetal death of unspecified cause	P95	Exclude	Only one case for 1997-2008		
151	Interpersonal violence unspecified means	Y09, Y36	141-142	Proportionately redistribute by age and sex to 141 and 142		1
152	Unspecified sexually transmitted disease	A64	3	Add to 3 (STDs, excluding HIV)	1	
153	Unspecified intestinal parasitism	B82	7	Add to 7 (Intestinal parasites, other parasites and vector diseases)	1	
154	Unspecified parasitic disease	B89	7	Add to 7 (Intestinal parasites, other parasites and vector diseases)	1	
155	Malignant neoplasm of other and ill-defined sites in the lip, oral cavity and pharynx	C14	28-29	Proportionally redistributed by age and sex to cancers of the oral cavity and pharynx. Denominator (28-29)	2	
156	Malignant neoplasm of other and ill-defined digestive organs	C26	30-35	Proportionally redistributed by age and sex to cancers of the oesophagus, stomach, colo-rectal, liver, gall bladder and pancreas. Denominator (30-35)	2	
157	Malignant neoplasm of other and ill-defined sites in the respiratory system and intrathoracic organs	C39	36-37	Proportionally redistributed by age and sex to cancers of the larynx and trachea. Denominator (36-37)	2	

NBD analysis codes	Single disease	ICD-10 codes	Target NBD analysis codes	Redistributions	Sequence for redistribution	
					Natural deaths	Injury deaths
158	Unspecified transport accident	V99	124-126	Proportional redistribution by age and sex to 124-126		2
159	Exposure to unspecified factor	X59	124-139	Proportional redistribution by age and sex to 124-139		6
160	Injuries with undetermined intent	Y10-Y34	124-142	Proportional redistribution by age and sex to 124-142		7
161	Septicaemia	A40-A41				
162	Sequence of hyper alimentation	E68	57		2	
163	Volume depletion	E86	143	Combine with 143 and proportionally redistributed by age and sex to all natural deaths. Denominator (1-123)	6	
164	Other disorders of fluid, electrolyte and acid-base balance	E87	143	Combine with 143 and proportionally redistributed by age and sex to all natural deaths. Denominator (1-123)	6	
165	Toxic encephalopathy	G92	127, 136	Proportional redistribution by age and sex to 127, 136		1
166	Pulmonary embolism	I26	143	Combine with 143 and proportionally redistributed by age and sex to all natural deaths. Denominator (1-123)	6	
167	Atrioventricular and left bundle-branch block	I44	87	Add to 87 (conduction disorders and other dysrhythmias)	1	
168	Other conduction disorders	I45	87	Add to 87 (Conduction disorders and other dysrhythmias)	1	
169	Arterial embolism and thrombosis	I74	89,90	Proportionally redistributed by age and sex to peripheral vascular disorders and other circulatory disorders. Denominator (89, 90)	2	
170	Portal vein thrombosis	I81	103	Add to 103 (other digestive)	1	
*171	Pneumonitis due to solids and liquids	J69	143	Combine with 143 and proportionally redistributed by age and sex to all natural deaths. Denominator (1-123)	6	
172	Adult respiratory distress syndrome	J80	143	Combine with 143 and proportionally redistributed by age and sex to all natural deaths. Denominator (1-123)	6	
173	Pulmonary oedema	J81	205; Redist 81-85, 87, 114	Combine with 205 and proportionately redistribute by age and sex to rheumatic heart disease, hypertensive heart disease, IHD, pericarditis, etc., cardiomyopathy, conduction disorders and other dysrhythmias, congenital heart anomalies. (Denominator (81--85, 87, 114 [<15years])	2	
174	Oesophageal varices	I85	100	Add to 100 (cirrhosis of liver)	1	
175	Pyothorax	J86	11	Add to 11 (lower respiratory infections)	1	
176	Pleural effusion, not elsewhere classified	J90	1	Add to 1 (tuberculosis)	1	
177	Pneumothorax	J93	143	Combine with 143 and proportionally redistributed by age and sex to all natural deaths. Denominator (1-123)	6	
178	Other pleural conditions	J94	1	Add to 1 (tuberculosis)	1	

NBD analysis codes	Single disease	ICD-10 codes	Target NBD analysis codes	Redistributions	Sequence for redistribution	
					Natural deaths	Injury deaths
179	Peritonitis	K65	15, 18, 96-98, 101,102	Proportionally redistributed by age and sex to maternal sepsis, abortion, peptic ulcer, appendicitis, intestinal obstructions and strangulated hernia, gall bladder and pancreatitis. Denominator (15, 18, 96-98, 101, 102)	2	
180	Other disorders of peritoneum	K66	15, 18, 96-98, 101,102	Proportionally redistributed by age and sex to maternal sepsis, abortion, peptic ulcer, appendicitis, intestinal obstructions and strangulated hernia, gall bladder and pancreatitis. Denominator (15, 18, 96-98, 101, 102)	2	
181	Toxic liver disease	K71	103	Add to 103 (other digestive)	1	
182	Hepatic failure, not elsewhere classified	K72	100	Add to 100 (cirrhosis of liver)	1	
183	Other inflammatory liver diseases	K75	100	Add to 100 (cirrhosis of liver)	6	
184	Haematemesis and melaena/other diseases of GIT	K92	31,32,96	Proportionally redistributed by age and sex to stomach, colorectal, peptic ulcer. Denominator (131, 32, 96)	2	
**185	Osteomyelitis	M86	143	Combine with 143 and proportionally redistributed by age and sex to all natural deaths. Denominator (1-123)	6	
186	Acute renal failure	N17	104	Add to 104 (acute and chronic glomerulo- and pyelonephritis)	1	
187	Chronic renal failure	N18	104	Add to 104 (acute and chronic glomerulo- and pyelonephritis)	1	
188	Unspecified renal failure	N19	104	Add to 104 (Acute and chronic glomerulo- and pyelonephritis)	1	
189	Gangrene, not elsewhere classified	R02	143	Combine with 143 and proportionally redistributed by age and sex to all natural deaths. Denominator (1-123)	6	
190	Ascites	R18	143	Combine with 143 and proportionally redistributed by age and sex to all natural deaths. Denominator (1-123)	6	
191	Sequelae of other and unspecified infectious and parasitic diseases	B94	10	Add to 10 (other infectious diseases)	1	
192	Sequelae of malnutrition and other nutritional deficiencies	E64	24-27	Proportionally redistributed by age and sex to protein energy malnutrition, vitamin-A deficiency, iron-deficiency anaemia and other nutritional disorders. Denominator (24-27)	2	
193	Sequelae of inflammatory diseases of central nervous system	G09	8	Add to 8 (meningitis and encephalitis)	1	
194	Cerebral palsy	G80 for >=20 years	75	Add to 75 other neurological if > =20 years	1	
195	Hemiplegia	G81	86	Add to 86 (stroke)	1	
196	Paraplegia and tetraplegia	G82	124-126,128	Proportional redistribution by age and sex to 124-126, 128	2	
197	Other paralytic syndromes	G83	124-126, 128	Proportional redistribution by age and sex to 124-126, 128	2	

NBD analysis codes	Single disease	ICD-10 codes	Target NBD analysis codes	Redistributions	Sequence for redistribution	
					Natural deaths	Injury deaths
198	Sequelae of transport accidents	Y85	158	Add to 158 and then proportional redistribution by age and sex to 124–126		1
199	Sequelae of other accidents	Y86	127-139	Proportionately redistribute by age and sex to 127-139		2
200	Sequelae of intentional self-harm, assault and events of undetermined intent	Y87	140-142, 160	Proportionately redistribute by age and sex 140-142, 160		2
201	Sequelae of other external causes	Y89	141-142; 159	Proportionately redistribute by age and sex to 141-142; 159		3
202	Disseminated intravascular coagulation [defibrination syndrome]	D65	143	Combine with 143 and proportionally redistributed by age and sex to all natural deaths. Denominator (1-123)	6	
203	Cardiac arrest	I46	143	Combine with 143 and proportionally redistributed by age and sex to all natural deaths. Denominator (1-123)	6	
204	Other cardiac arrhythmias	I49	81-85; 87, 90	Proportionally redistributed by age and sex to rheumatic heart disease, hypertensive heart disease, IHD, pericarditis, etc., cardiomyopathy, conduction disorders and other dysrhythmias and other circulatory disease. Denominator (81-85 + 87 + 90)	2	
205	Heart failure	I50	Add to 173, Redist to 81-85, 87, 114	Combine with 173 and proportionately redistribute by age and sex to rheumatic heart disease, hypertensive heart disease, IHD, pericarditis, etc., cardiomyopathy, conduction disorders and other dysrhythmias, congenital heart anomalies. (Denominator (81-85, 87, 114 [<15 years only]))	5	
206	Respiratory failure, not elsewhere classified	J96	91-95	Proportionally redistributed by age and sex to COPD, pneumoconiosis, asthma, other interstitial lung disease, other respiratory. Denominator (91-95)	2	
207	Secondary hypertension	I15	143	Combine with 143 and proportionally redistributed by age and sex to all natural deaths. Denominator (1-123)	1	
208	Atherosclerosis	I70	83, 86, 88, 89	Proportionally by age and sex to IHD, stroke, aortic aneurism and peripheral vascular disease. Denominator (83 + 86 + 88 + 89)	3	
209	Other disorders of brain (encephalitis)	G93	8	Add to 8 (meningitis and encephalitis)	1	

* Ideally across injuries as well to allow for poisonings.

**Ideally to injuries as well but this was not done in this study.

4.6 Mid-year population estimates

We decided to use the alternative mid-year population estimates produced by Dorrington (2013) for this research in preference to the official mid-year estimates for several reasons, in particular, because these estimates were constructed to have an age distribution consistent

with that of the 2001 and 2011 census populations, and are available in the age categories needed.

These estimates were derived by projecting the 2011 census population (by single years of age, sex, population group and province) backwards to 2001 using survival factors from the ASSA2008 full model, taking into account in- and out-migration and immigration as estimated from answers to questions on migration in the 2011 Census and 2007 Community Survey. Emigration was set in total to match the numbers in the official mid-year estimates (Stats SA, 2013b).

As the alternative mid-year estimates were only available for the years 2001-2013, we derived estimates for the years prior to 2001 by continuing the back-projection process to 1996.

5 Data analysis

5.1 Age-standardised rates

The age-specific mortality rates have been calculated using the mid-year population estimates described in section 4.6. These were then used to calculate age-standardised rates by applying the WHO world population standard (Ahmad *et al*, s.a.).

5.2 Years of life lost

Two methods were used to calculate the years of life lost (YLLs), namely:

a) YLLs have been calculated using the same method applied in the 2010 Global Burden of Disease Study (Murray *et al*, 2012). In particular, no age weighting and no discounting were used, together with the newly created standard life expectancies.

b) YLLs have also been calculated using the same method used in the earlier Global Burden of Disease studies (Murray *et al*, 2001; WHO, 2003). The same standard life expectancy and a discount rate of 3% p.a. without age weighting have been applied.

$$YLL = \frac{N \times (1 - e^{-0.03L})}{0.03}$$

N = the number of deaths for a given age

L = the standard life expectancy at age of death for a given age

Effectively, the earlier standard life expectancies can be represented by a model life table, Coale and Demeny West Level 26, with a life expectancy at birth of 82.5 years for females (Coale and Demeny, 1966; Coale and Guang, 1989). An arbitrary biological male/female difference of life expectancy at birth of 2.5 years was used. This standard has a life expectancy at birth for males of 80 years, modelled on the West Level 25 life table for females. The discount rate of 3% per YLL is applied, and implies that individuals prefer time lived now to time in the future. This discount rate is recommended by the International Panel on Cost Effectiveness in Health and Medicine (Gold *et al*, 1996).

5.3 Chronic and acute care needs

The causes of death in the NBD list were categorised according to the broad care needs, distinguishing causes that would require acute and chronic care as proposed by Setel *et al* (2004). For 20 of the causes, it was difficult to establish whether these would have required chronic or acute care, hence these remained 'uncategorised' (Table 7).

Table 6: SA NBD list causes categorised according to chronic versus acute care needs

Chronic care		Acute care		Uncategorised	
ZA 1	Tuberculosis	ZA 3	Sexually transmitted infections excluding HIV	ZA 106	Renal disease
ZA 2	HIV/AIDS	ZA 4	Diarrhoeal diseases	ZA 11	Other infectious diseases
ZA 7	Hepatitis	ZA 5	Vaccine-preventable childhood	ZA 14	Otitis media
ZA 17	Hypertensive disorders of pregnancy	ZA 6	Meningitis and encephalitis	ZA 70	Other mental and behavioural disorders
ZA 26	Protein-energy malnutrition	ZA 9	Malaria	ZA 76	Non-migraine headache
ZA 27	Vitamin-A deficiency	ZA 10	Intestinal parasites, other parasitic and vector diseases	ZA 77	Other neurological conditions
ZA 28	Iron-deficiency anaemia	ZA 12	Lower respiratory infections	ZA 81	Other vision loss
ZA 29	Other nutritional disorders	ZA 13	Upper respiratory infections	ZA 82	Other sense organ disorders
ZA 30	Mouth cancer	ZA 15	Maternal haemorrhage	ZA 96	Other interstitial lung disease
ZA 31	Oropharynx, nasopharynx and other pharynx cancers	ZA 16	Maternal sepsis	ZA 97	Other respiratory
ZA 32	Oesophageal cancer	ZA 18	Obstructed labour	ZA 105	Other digestive diseases
ZA 33	Stomach cancer	ZA 19	Abortion	ZA 108	Other urinary and gynaecological diseases
ZA 34	Colo-rectal cancer	ZA 21	Other maternal	ZA 109	Skin diseases
ZA 35	Liver cancer	ZA 22	Preterm birth complications	ZA 112	Back and neck pain
ZA 36	Gallbladder and biliary tract cancers	ZA 23	Birth asphyxia and trauma	ZA 113	Other musculo-skeletal disorders
ZA 37	Pancreas cancer	ZA 24	Sepsis and other new born infectious conditions	ZA 114	Neural tube defects
ZA 38	Larynx cancer	ZA 25	Other non-infectious conditions arising in the perinatal period	ZA 116	Congenital heart anomalies
ZA 39	Trachea, bronchi and lung cancers	ZA 99	Appendicitis	ZA 117	Congenital disorders of GIT
ZA 40	Melanoma cancer	ZA 100	Intestinal obstruction and strangulated hernia	ZA 118	Urogenital malformations
ZA 41	Other skin cancers	ZA 101	Non-infective inflammatory bowel disease	ZA 119	Fetal alcohol syndrome
ZA 42	Breast cancer	ZA 126	Road injuries	ZA 121	Other chromosomal abnormalities
ZA 43	Cervix cancer	ZA 127	Other transport accidents	ZA 122	Other congenital

Chronic care		Acute care		Uncategorised	
					abnormalities
ZA 44	Corpus uteri cancer	ZA 128	Poisonings (including herbal)	ZA 125	Other oral diseases
ZA 45	Ovary cancer	ZA 129	Falls	ZA 132	Hanging, strangulation and other threats to breathing
ZA 46	Prostate cancer	ZA 130	Fires, heat and hot substances		
ZA 47	Testis cancer	ZA 131	Drowning		
ZA 48	Bladder cancer	ZA 133	Mechanical forces		
ZA 49	Kidney cancer	ZA 134	Exposure to natural forces		
ZA 50	Brain cancer	ZA 135	Adverse effects of medical and surgical treatment		
ZA 51	Thyroid cancer	ZA 136	Animal contact		
ZA 52	Hodgkin's lymphoma	ZA 137	Other unintentional injuries		
ZA 53	Non-Hodgkin's lymphoma	ZA 138	Self-inflicted injuries		
ZA 54	Multiple myeloma	ZA 139	Interpersonal violence		
ZA 55	Leukaemia	ZA 8	Septicaemia		
ZA 56	Other malignant neoplasms				
ZA 57	Benign neoplasms				
ZA 58	Diabetes mellitus				
ZA 59	Endocrine nutritional, blood and immune disorders				
ZA 60	Unipolar depressive disorder				
ZA 61	Bipolar affective disorder				
ZA 62	Schizophrenia				
ZA 63	Alcohol dependence				
ZA 64	Drug use				
ZA 65	Anxiety disorders				
ZA 66	Eating disorders				
ZA 67	Development disorders				
ZA 68	Childhood behaviour disorders				
ZA 69	Mental retardation not included as sequelae elsewhere				
ZA 71	Alzheimer's and other degenerative diseases of				

Chronic care		Acute care		Uncategorised	
	the nervous system				
ZA 72	Parkinson's disease				
ZA 73	Multiple sclerosis				
ZA 74	Epilepsy				
ZA 75	Migraine				
ZA 78	Glaucoma				
ZA 79	Cataracts				
ZA 80	Hearing loss not due to other diseases or injuries				
ZA 83	Rheumatic heart disease				
ZA 84	Hypertensive heart disease				
ZA 85	Ischaemic heart disease				
ZA 86	Pericarditis, endocarditis and myocarditis				
ZA 87	Cardiomyopathy				
ZA 88	Cerebrovascular disease				
ZA 89	Conduction disorders and other dysrhythmias				
ZA 90	Aortic aneurism				
ZA 91	Peripheral vascular disorders				
ZA 92	Other circulatory diseases				
ZA 93	Chronic obstructive pulmonary disease				
ZA 94	Pneumoconiosis				
ZA 95	Asthma				
ZA 98	Peptic ulcer disease				
ZA 102	Cirrhosis and liver disease				
ZA 103	Gall bladder and bile duct disease				
ZA 104	Pancreatitis				
ZA 107	Benign prostatic hypertrophy				
ZA 110	Rheumatoid arthritis				
ZA 111	Osteoarthritis				
ZA 115	Cleft lip and cleft palate				

Chronic care		Acute care		Uncategorised	
ZA 120	Down syndrome				
ZA 121	Other chromosomal abnormalities				
ZA 123	Dental caries				
ZA 124	Periodontal disease				

References

Ahmad OB, Boschi-Pinto C, Lopez AD, Murray CJL, Lozano R, Inoue M. s.a. *Age Standardisation of Rates: A new WHO standard*. (GPE Discussion Paper Series No 31). Geneva: EIP/GPE/EBD WHO. Available at http://www.who.int/healthinfo/paper31.pdf [Accessed, 10 February 2010].

Black RE, Cousens S, Johnson HL, Lawn JE, Rudan I, Bassani DG, Jha P, Campbell H, Walker CF, Cibulskis R, Eisele T, Liu L, Mathers C; Child Health Epidemiology Reference Group of WHO and UNICEF. Global, regional, and national causes of child mortality in 2008: a systematic analysis. *Lancet.* 2010; **375**(9730):1969-1987. Epub 2010 May 11.

Bradshaw D, Dorrington RE, Laubscher R. *Rapid mortality surveillance report 2011.* Cape Town: South African Medical Research Council, 2012. ISBN: 978-1-920618-00-1.

Bradshaw D, Groenewald P, Laubscher R, Nannan N, Nojilana B, Norman R, Pieterse D, Schneider M, Bourne D, Timæus I, Dorrington RE, Johnson L. Initial burden of disease estimates for South Africa, 2000. *SAMJ* 2003; **93**:682-688.

Bradshaw D, Nannan N, Groenewald P, Joubert J, Laubscher R, Nojilana B, Norman R, Pieterse D and Schneider M. Provincial mortality in South Africa, 2000: priority-setting for now and a benchmark for the future. *SAMJ* 2005*;* **95**(7):496-503.

Bradshaw D, Pillay-vanWyk V, Laubscher R, Nojilana B, Groenewald P, Nannan N. *Cause of death statistics for South Africa: Challenges and possibilities for improvement*. Cape Town: Medical Research Council, 2011.

Coale AJ, Guang G. Revised regional model life tables at very low levels of mortality. *Population Index* 1989; **55**(4):613-643.

Coale AJ, Demeny P. *Regional Model Life Tables and Stable Population*. USA: Princeton University Press, 1966.

Chopra M, Lawn JE, Sanders DE, Barron P, Abdool Karim SS, Bradshaw D, Jewkes R, Abdool Karim Q, Flisher AJ, Mayosi BM, Tollman SM, Churchyard GJ, Coovadia H, for The Lancet South Africa team. Health in South Africa 6. Achieving the health Millennium Development Goals for South Africa: challenges and priorities. *Lancet* 2009; **374**(9694):1023-1031. Epub 2009 Aug 24.

Darikwa TB. *Estimating the level and trends of child mortality in South Africa, 1996-2006.* Thesis submitted in partial fulfilment of the Degree of Master of Philosophy in the Faculty of Commerce University of Cape Town, 2009.

Darikwa TB, Dorrington RE. The level and trends of child mortality in South Africa, 1996-2006. *Journal of African Population Studies* 2011;**25**,1 (Supplement):158-172.

Dorrington R, Bradshaw D, Laubscher R. *Rapid Mortality Surveillance 2012*. Cape Town: South African Medical Research Council, 2014. ISBN: 978-1-920618-19-3. Available at http://www.mrc.ac.za/bod/RapidMortalitySurveillanceReport2012.pdf

Dorrington RE. 2013. *Alternative South African mid-year estimates, 2013*. Centre for Actuarial Research Monograph 13, University of Cape Town, 2014. ISBN: 978-0-7992-2501-3. Available at https://www.commerce.uct.ac.za/Research_Units/CARE/Monographs/Monographs/Mono13.pdf

Dorrington RE, Moultrie TA, Timæus I. *Estimation of mortality using the South African Census 2001 data. Cape Tow*n: Centre for Actuarial Studies, University of Cape Town, 2004. ISBN: 0-7992-2281-X. Available at https://www.commerce.uct.ac.za/Research_Units/CARE/Monographs/Monographs/Mono11.pdf

GBD 2005. *Operations Manual*. Final Draft. Harvard University, Institute for Health Metrics and Evaluation at the University of Washington, John Hopkins University, Queensland University and the WHO, 2009. Available at http://www.globalburden.org/GBD_Study_Operations_Manual_Jan_20_2009.pdf

Gold MR, Siegal JE, Russell LG, *et al*. *Cost effectiveness in Health and Medicine*. Oxford: Oxford University Press, 1996.

Lawn JE, Kerber K, Enweronu-Laryea C, Cousens S. 3.6 million neonatal deaths--what is progressing and what is not? *Semin Perinatol* 2010 ;**34**(6):371-386.

Mathews S, Abrahams N, Jewkes R, Martin LJ, Lombard C. *Child homicide patterns in South Africa: Is there a link to child abuse?* Cape Town: South African Medical Research Council Research Brief, 2012.

Matzopoulos R, Prinsloo M, Bradshaw D, Pillay-van Wyk V, Gwebushe N, Mathews S, Martin L, Laubscher R, Lombard C, Abrahams N. *The Injury Mortality Survey: A national study of injury mortality levels and causes in South Africa in 2009.* Cape Town: South African Medical Research Council, 2013.

Murray CJL, Lopez AD. *The Global Burden of Disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020*. Cambridge MA: Harvard University Press on behalf of the World Health Organization and The World Bank, 1996.

Murray CJL, Ezzati M, Flaxman AD, *et al*. GBD 2010: design, definitions, and metrics. *Lancet* 2012;**380**:2063–2066.

Naghavi M, Makela S, Foreman K, O'Brien J, Pourmalek F, Lozano R. Algorithms for enhancing public health utility of national causes-of-death data. *Popul Health Metr* 2010; **8**:9.

Nannan N, Dorrington RE, Laubscher R, Zinyakatira N, Prinsloo M, Darikwa TB, Matzopoulos R, Bradshaw D. *Under-5 mortality statistics in South Africa: Shedding some light on the trends and causes 1997-2007.* Technical Report. Cape Town: South African Medical Research Council, 2012. Available at http://www.mrc.ac.za/bod/MortalityStatisticsSA.pdf

Nojilana B, Groenewald P, Bradshaw D, Reagon G. Quality of cause of death certification at an academic hospital in Cape Town, South Africa. *SAM J* 2009;**99**(9):648-652.

Nojilana B, Brewer L, Bradshaw D, Groenewald P, Burger EH, Levitt NS. Certification of diabetes-related mortality: the need for an international guideline. *J Clin Epidemiol* 2013;**66**(2):236-237.

Norman R, Bradshaw D, Schneider M, Joubert J, Groenewald P, Lewin S, Steyn K, Vos T, Laubscher R, Nannan N, Nojilana B, Pieterse D; South African Comparative Risk Assessment Collaborating Group. A comparative risk assessment for South Africa in 2000: towards promoting health and preventing disease. *SAM J* 2007; **97**(8 Pt 2):637-641.

Setel PW, Saker L, Unwin NC, Hemed Y, Whiting DR, Kitange H. Is it time to reassess the categorization of disease burdens in low-income countries? *Am J Public Health* 2004;**94**(3): 384-388.

Statistics South Africa (Stats SA). *Community Survey 2007.* Statistical Release P0301. Pretoria: Statistics South Africa, 2007. Available at http://www.statssa.gov.za/publications/P0301/P0301.pdf

Statistics South Africa (Stats SA). *Mortality and causes of death in South Africa, 2006: Findings from death notification.* Statistical Release P0309.3. Pretoria: Statistics South Africa, 2008.

Statistics South Africa (Stats SA). *Mortality and causes of death in South Africa, 2010: Findings from death notification*. Statistical Release P0309.3. Pretoria: Statistics South Africa, 2013a.

Statistics South Africa (Stats SA). *Mid-year population estimates, 2013*. Pretoria: Statistics South Africa, 2013b.

Statistics South Africa (Stats SA). *Census 2011 Statistical Release (revised) – P0301.4*. Pretoria: Statistics South Africa, 2013c.

Statistics South Africa (Stats SA). *Census 2011: 10% Sample of unit records*. DVD issued by Statistics South Africa, April 2014, 2014a.

Statistics South Africa (Stats SA). *An assessment of household deaths collected during Census 2011: A discussion document.* Statistical release D0301.5. Pretoria: Statistics South Africa, 2014b.

Statistics South Africa (Stats SA). *Mortality and causes of death in South Africa, 2011: Findings from death notification*. Statistical Release P0309.3. Pretoria: Statistics South Africa, 2014c.

World Health Organization. *International Statistical Classification of Disease and Related Health Problems, Vol. 2*. 10th revision. Geneva: World Health Organization; 2008.

World Health Organization. Analysing mortality levels and cause-of-death (ANACoD) electronic tool. Geneva: World Health Organization Department of Health Statistics and Health Information, 2011.

Appendices

Appendix 1: Mapping NBD analysis codes, ICD-10 codes and ZA codes

Table A1: Mapping NBD analysis codes, ICD-10 codes and ZA codes

NBD analysis codes	Single disease	ICD-10 codes	Target ZA code
1	Tuberculosis	A15-A16, B90, U51, U52	1
2	HIV/AIDS	B20-B24	2
3	Sexually transmitted infections excluding HIV	A50>=1 month, A51-A53, A54>=1 month, A55-A63, N70-N73	3
4	Diarrhoeal diseases	A00-A09, K52	4
5	Selected vaccine-preventable childhood diseases	A33-A35>=1month, A36-A37, A80, B01, B05-B06, B26, B91	5
6	Malaria	B50-B54	9
7	Intestinal parasites, other parasitic and vector diseases	A82, A90-A91, A95, B55-B57, B65, B67-B74, B76-B81	10
8	Meningitis and encephalitis	A39, A87, G00-G03, G05	6
9	Hepatitis	B15-B19	7
10	Other infectious diseases	A20-A28, A30-A32, A38, A42-A49, A65-A69, A70-A74, A75-A79, A81, A88-A89, A92vA94, A96-A99, B00, B02-B04, B07-B09, B25, B27-B32, B34, B37-B44, B46-B49, B58, B60-B64, B66, B75, B83, B92, B95-B99	11
11	Lower respiratory infections	J09-J18, J20-J22,P23 >6 days	12
12	Upper respiratory infections	J00-J06	13
13	Otitis media	H65-H66	14
14	Maternal haemorrhage	O20, O44-O46, O67, O72	15
15	Maternal sepsis	085-086	16
16	Hypertensive disorders of pregnancy	O10-O16	17
17	Obstructed labour	O64-O66	18
18	Abortion	O00-O08	19
19	Other maternal	O21-O43, O47-O48, O60-O63, O68-O71, O73-O75, O80-O84, O87-O92, O94-O97	21
20	Pre-term birth complications	P01, P07, P22, P23(<6days),P25-P28, P29, P52, P61, P77, P80	22
21	Birth asphyxia and trauma	P02, P03, P10-P15, P20-21, P24, P50, P90, P91	23
22	Sepsis and other newborn infectious conditions	P35, P36-P39, P60, (A33-A35, A50,A54) <1month	24
23	Other non-infectious conditions arising in the perinatal period	P00, P04, P05, P08, P51, P53-P59, P62-P74, P76, P78-P79, P81-P89, P93-P94	25
24	Protein-energy malnutrition	E40-E46	26
25	Vitamin-A deficiency	E50	27
26	Iron-deficiency anaemia	D50	28
27	Other nutritional disorders	D51-D53, E00-E02, E51-E63	29
28	Mouth cancer	C00-C08	30
29	Oropharynx, nasopharynx and other pharynx cancers	C09-C13	31
30	Oesophageal cancer	C15	32

NBD analysis codes	Single disease	ICD-10 codes	Target ZA code
31	Stomach cancer	C16	33
32	Colo-rectal cancer	C18-C21	34
33	Liver cancer	C22	35
34	Gallbladder and biliary tract cancers	C23-C24	36
35	Pancreas cancer	C25	37
36	Larynx cancer	C32	38
37	Trachea, bronchi and lung cancers	C33-C34	39
38	Melanoma cancer	C43	40
39	Other skin cancers	C44	41
40	Breast cancer	C50	42
41	Cervix cancer	C53	43
42	Corpus uteri cancer	C54-C55	44
43	Ovary cancer	C56	45
44	Prostate cancer	C61	46
45	Testis cancer	C62	47
46	Bladder cancer	C67	48
47	Kidney cancer	C64-C66, C68	49
48	Brain cancer	C70-C72	50
49	Thyroid cancer	C73	51
50	Hodgkin's lymphoma	C81	52
51	Non-Hodgkin's lymphoma	C82-C85, C96	53
52	Multiple myeloma	C88, C90	54
53	Leukaemia	C91-C95	55
54	Other malignant neoplasms	C17, C30-C31, C37-C38, C40-C41, C45, C47-C49, C51-C52, C57-C58, C60, C63, C69, C74-C75, C77-C79, C97	56
55	Benign neoplasms	D01-D48	57
56	Diabetes mellitus	E10-E14	58
57	Endocrine nutritional, blood and immune disorders	D55-D62, D64, D66-D83, D86-D89,E03-E07, E15-E16, E20-E32, E34, E65-E67, E70-E80, E83-E85, E88	59
58	Unipolar depressive disorder	F32-F33	60
59	Bipolar affective disorder	F30-F31	61
60	Schizophrenia	F20-F29	62
61	Alcohol dependence	F10	63
62	Drug use	F11-F16, F18-F19, G13	64
63	Anxiety disorders	F40-F44	65
64	Eating disorders	F50	66
65	Development disorders	F80-F82, F84, F88-F89	67
66	Childhood behaviour disorders	F90-F92	68
67	Mental retardation not included as sequelae elsewhere	F70-F73, F78, F79	69
68	Other mental and behavioural disorders	F04-F07, F09, F17, F34-F39, F45-F48, F51-F59, F60-F69, F93-F98	70
69	Alzheimer's and other degenerative diseases of the nervous system	F01, F03, G30-G31	71

NBD analysis codes	Single disease	ICD-10 codes	Target ZA code
70	Parkinson's disease	G20-G21	72
71	Multiple sclerosis	G35	73
72	Epilepsy	G40-G41	74
73	Migraine	G43	75
74	Non-migraine headache	G44	76
75	Other neurological conditions	G06-G08, G10-G12, G14, G22-G25, G32, G36-37, G45-G47, G50-G62, G64, G70-G72, G90-G91, G95-G98	77
76	Glaucoma	H40	78
77	Cataracts	H25-H26	79
78	Hearing loss not due to other diseases or injuries	H90-H91	80
79	Other vision loss	H30-H35, H49-H54	81
80	Other sense organ disorders	H00-H21, H27, H43-H47, H55-H57, H60-H61, H68-H83, H92-H93	82
81	Rheumatic heart disease	I00-I09	83
82	Hypertensive heart disease	I11	84
83	Ischaemic heart disease	I20-I25	85
84	Pericarditis, endocarditis and myocarditis	I30, I31, I33, I38-I40	86
85	Cardiomyopathy	I42	87
86	Cerebrovascular disease	I60-I69	88
87	Conduction disorders and other dysrhythmias	I47-I48	89
88	Aortic aneurism	I71	90
89	Peripheral vascular disorders	I73	91
90	Other circulatory diseases	I27, I28, I34-I37, I72, I77-I78, I80, I82-I84, I86-I89, I95-I98	92
91	Chronic obstructive pulmonary disease	J40-J44	93
92	Pneumoconiosis	J60-J65	94
93	Asthma	J45-J46	95
94	Other interstitial lung disease	J84	96
95	Other respiratory	J30-J39, J47, J66-J68, J70, J82, J85, J91 J92, J95, J98-J99	97
96	Peptic ulcer disease	K25-K27	98
97	Appendicitis	K35-K37	99
98	Intestinal obstruction and strangulated hernia	K40-K46, K56	100
99	Non-infective inflammatory bowel disease	K50-K51	101
100	Cirrhosis and liver disease	K70, K73, K74	102
101	Gall bladder and bile duct disease	K80-K83	103
102	Pancreatitis	K85, K86	104
103	Other digestive diseases	K20-K22, K28-K31, K38, K55, K57-K63, K76, K77, K90-K91	105
104	Renal disease	N00-N08, N10-N12, I12, I13	106
105	Benign prostatic hypertrophy	N40	107

NBD analysis codes	Single disease	ICD-10 codes	Target ZA code
106	Other urinary and gynaecological diseases	N13-N15, N20-N21, N23, N25-N39, N41-N50, N60-N64, N75-N96, N97-N98	108
107	Skin diseases	L00-L08, L10-L13, L20-L30, L40-L98, B35-B36, B85-B88	109
108	Rheumatoid arthritis	M05-M06	110
109	Osteoarthritis	M15-M19	111
110	Back and neck pain	M47-M54	112
111	Other musculoskeletal disorders	M00-M02, M08, M10-M13, M20-M46, M60-M85, M87-M99	113
112	Neural tube defects	Q00, Q03, Q05	114
113	Cleft lip and cleft palate	Q35-Q37	115
114	Congenital heart anomalies	Q20-Q28	116
115	Congenital disorders of GIT	Q38-Q45	117
116	Urogenital malformations	Q50-Q64	118
117	Fetal alcohol syndrome	Q86	119
118	Down syndrome	Q90	120
119	Other chromosomal abnormalities	Q91-Q99	121
120	Other congenital abnormalities	Q01, Q02, Q04, Q06, Q07, Q10-Q18, Q30-Q34, Q65-Q85, Q87-Q89, P75	122
121	Dental caries	K02	123
122	Periodontal disease	K05	124
123	Other oral diseases	K00-K01, K03-K04, K06-K14	125
124	Road injuries	V01-V04,V06, V09-V80, V82-V85, V87, V89	126
125	Rail injury	V05, V81	127
126	Other transport accidents	V86, V88, V91, V93-V98	127
127	Poisonings (including herbal)	X40-X49, Y67	128
128	Falls	W00-W19	129
129	Fires, heat and hot substances	X00-X19	130
130	Drowning	V90, V92, W65-W70, W73, W74	131
131	Other threats to breathing	W75-W76, W78-W84	132
132	Mechanical forces: machinery	W24, W30-W31	133
133	Mechanical forces; firearm	W32-W34	133
134	Mechanical forces: sharp object	W25-W29, W45-W46	133
135	Exposure to natural forces	X30-X39	134
136	Adverse effects of medical and surgical treatment	Y39-Y66, Y68-Y84, Y88	135
137	Mining accidents	Y37, W77	132
138	Animal contact	W53-W59, X20-X27, X29	136
139	Other unintentional injuries	W20-W23, W35-W44, W49-W52, W60, W64, W85-W94, W99, X28, X50-X58, Y38	137
140	Self-inflicted injuries	X60-X84	138
141	Interpersonal violence with firearm	X93-X95	139
142	Interpersonal violence without firearm	X85-X92, X96-X99, Y00-Y08	139
143	Ill-defined natural: redistribute to all causes	R00-R01, R03-R17, R19-R23, R25-R54, R56, R59-R63, R68, R70-R99	1-125
144	Ill-defined: redistribute tosome causes	R55, R57, R58, R64	1-125

NBD analysis codes	Single disease	ICD-10 codes	Target ZA code
145	Malignant neoplasm without specification of site	C80, C76	30-56
146	Ill-defined descriptions of heart disease and unspecified disorders of circulatory system	I51, I99	83-87, 89, 92
147	Hypertension only	I10	86, 88, 104
148	Mental disorder, not otherwise specified	F99	60-70
149	Perinatal death unspecified cause	P96, P92, G80 <20 years	22-25
150	Fetal death of unspecified cause	P95	excluded
151	Interpersonal violence unspecified means	Y09, Y36	139
152	Unspecified sexually transmitted disease	A64	3
153	Unspecified intestinal parasitism	B82	10
154	Unspecified parasitic disease	B89	10
155	Malignant neoplasm of other and ill-defined sites in the lip, oral cavity and pharynx	C14	30-31
156	Malignant neoplasm of other and ill-defined digestive organs	C26	32-37
157	Malignant neoplasm of other and ill-defined sites in the respiratory system and intra-thoracic organs	C39	38-39
158	Unspecified transport accident	V99	126-127
159	Exposure to unspecified factor	X59	126
160	Injuries with undetermined intent	Y10-Y34	126-130
161	Septicaemia	A40-A41	8
162	Sequence of hyper alimentation	E68	59
163	Volume depletion	E86	1-125
164	Other disorders of fluid, electrolyte and acid-base balance	E87	1-125
165	Toxic encephalopathy	G92	128,135
166	Pulmonary embolism	I26	1-125
167	Atrioventricular and left bundle-branch block	I44	89
168	Other conduction disorders	I45	89
169	Arterial embolism and thrombosis	I74	91, 92
170	Portal vein thrombosis	I81	105
171	Pneumonitis due to solids and liquids	J69	1-125
172	Adult respiratory distress syndrome	J80	1-125
173	Pulmonary oedema	J81	83-87, 89,116
174	Oesophageal varices	I85	102
175	Pyothorax	J86	12
176	Pleural effusion, not elsewhere classified	J90	1
177	Pneumothorax	J93	1-125
178	Other pleural conditions	J94	1
179	Peritonitis	K65	16, 19, 98-100, 103, 104

NBD analysis codes	Single disease	ICD-10 codes	Target ZA code
180	Other disorders of peritoneum	K66	16, 19, 98-100, 103, 104
181	Toxic liver disease	K71	105
182	Hepatic failure, not elsewhere classified	K72	102
183	Other inflammatory liver diseases	K75	100
184	Haematemesis and melaena/other diseases of GIT	K92	33, 34, 98
185	Osteomyelitis	M86	1-125
186	Acute renal failure	N17	106
187	Chronic renal failure	N18	106
188	Unspecified renal failure	N19	106
189	Gangrene, not elsewhere classified	R02	1-125
190	Ascites	R18	1-125
191	Sequelae of other and unspecified infectious and parasitic diseases	B94	11
192	Sequelae of malnutrition and other nutritional deficiencies	E64	26-29
193	Sequelae of inflammatory diseases of central nervous system	G09	6
194	Cerebral palsy	G80 for >=20 years	77
195	Hemiplegia	G81	88
196	Paraplegia and tetraplegia	G82	126-127, 129
197	Other paralytic syndromes	G83	126-127, 129
198	Sequelae of transport accidents	Y85	126-127
199	Sequelae of other accidents	Y86	128-137
200	Sequelae of intentional self-harm, assault and events of undetermined intent	Y87	126-130, 138-139
201	Sequelae of other external causes	Y89	126, 139
202	Disseminated intravascular coagulation (defibrination syndrome)	D65	1-125
203	Cardiac arrest	I46	1-125
204	Other cardiac arrhythmias	I49	83-87, 89, 92
205	Heart failure	I50	83-87, 89,116
206	Respiratory failure, not elsewhere classified	J96	93-95
207	Secondary hypertension	I15	1-125
208	Atherosclerosis	I70	85, 88, 90, 91
209	Other disorders of brain (encephalitis)	G93	6
210	Extra pulmonary TB	A17-A18	1
211	Miliary TB	A19	1
212	Encephalitis	A83-A86, G04	6
213	Indirect maternal	O98-099	20
214	Legal intervention and acts of war	Y35	140
215	HIV pseudonyms	B33, B45, B59, C46, D84	2

Appendix 2: Method of deriving completeness

Completeness was estimated prior to the release of the 10% unit record sample of the census (Stats SA, 2014a), the re-editing of the deaths reported by households (Stats SA, 2014b) and the 2011 cause-of-death report (Stats SA, 2014c). However, since the bulk of the editing of deaths reported by households appears to have been of the records with unstated ages and/or sex, and these were ignored for the purposes of this research, and since the approach described below involves smoothing and rescaling to match the national estimates, the new data can be expected to have only a limited impact on the estimates of completeness.

It is estimated that had the corrected census data been used, infant mortality (and hence also under-5 mortality) would have increased by 3 per mille in the Western Cape and reduced by 2 per mille in the Free State, and that the change would not have been more than 1 per mille for the other provinces. As far as adult mortality is concerned, the impact is likely to be negligible except for about a 1% increase in rates in Limpopo.

2.1 Estimates of infant and child mortality rates

Crude estimates of q0 and 5q0 for each province and the country as a whole were derived for 1996, 2001, 2006 and 2011. The estimates for 2001 are essentially those from Dorrington, Moultrie and Timæus (2004) with two provincial exceptions: for KwaZulu-Natal, q0 was estimated as the average of the estimate directly from deaths reported by households in the census4 and that estimated by Dorrington et al (2004) and 5q0 was estimated from q0 in the same ratio as estimated by Dorrington et al; and for North West q0 estimated directly from deaths reported by households in the census, since this produced a more sensible ratio to the estimate of 5q0 from Dorrington et al. The estimates for 2006 were derived by Darikwa applying the same method as outlined by Darikwa (2009) to each of the provinces. The estimates for 2011 were derived from the deaths reported by households in the census (Stats SA, 2013c) assuming that none of those with unspecified age (about 25% of total deaths) were childhood deaths, because including even a small proportion of these deaths resulted in estimates that appeared to be implausibly high. In the case of 1996, estimates from ASSA2008 (which was calibrated to several empirical sources of data including birth history

[4] $D_0/(P_0+0.75D_0)$, where D_0 represents deaths under the age of 1 reported by households and P_0 represents the census population under the age of 1.

data from the 1996 census and the 1998 SADHS) were used because the 1996 census did not include questions on deaths in households and therefore it was not possible to provide estimates from this source.

In order to improve the robustness of the estimates (i.e. smooth the estimates), we regressed (for q_0 and ${}_5q_0$ separately) the estimates for the provinces for each set (census/survey) on each other set using orthogonal regression. The estimate for a particular census/survey (provinces and national) was set to the average of the original estimate and the estimates implied by the regressions on the three other sets of estimates. In effect, this assumes that there is a linear relationship by level between the provincial estimates of any two sets of census/survey estimates.

2.2 Estimates of number of births

The numbers of births were estimated by projecting backwards the numbers of survivors counted at each age at the time of the 2011 Census using survival factors from the ASSA2008 lite model. For example, the back-projected numbers of those aged 0 last birthday at the census represent the number born within 1 year before the census, the back-projected numbers of those aged 1 last birthday at the census represent the survivors of the number born between 1 and 2 years before the census, etc. The numbers of births for each calendar year, up to and including 2010, were then interpolated from these numbers of births by census-year (i.e. the years ending on the anniversary of the census). In the case of the provinces, the numbers that were projected backwards were the numbers identified as being born in the province (and not those currently living in the province) to allow for interprovincial migration between birth and time of the survey.

2.3 Estimates of completeness of infant and under-5 deaths

Completeness of reporting of deaths under the age of 1 and under the age of 5 was estimated in 1996, 2001, 2006 and 2011 as the ratio of reported deaths to those expected on the basis of the number of births in the year, multiplied by the robust (smoothed) estimates of q_0 and ${}_5q_0$ described in Appendix 2.1. The estimates of completeness for 1997 were derived in a similar way, but using as estimates of q_0 and ${}_5q_0$ those derived by linear interpolation between the estimates for 1996 and 2001. Similarly, estimates of the completeness for 2010 were based on estimates derived from interpolating between the estimates for 2006 and 2011.

There was a sharp increase in the number of registered deaths from 1997 to 1998, nationally and in many of the provinces, which was more likely to be due to an increase in completeness than to an increase in mortality (see Section 4.1). Thus, the estimates of completeness for 1998 were derived by simply scaling up the estimate of completeness in 1997 by the ratio of registered deaths in 1998 to those in 1997.

Estimates of completeness for the remaining years were first derived by linear interpolation between the values for 1998, 2001, 2006 and 2010, with the exception of the completeness of reporting of deaths for 2006 to 2010 at the national level. Linear interpolation of completeness under the age of 1 for the period 2006 to 2010 would have implied a decline in the percentage of deaths registered from 88% to 76% (which may imply that the estimates of the number of births in this period are too high), while the proportion of births being registered up to the end of the calendar year after the year of birth averaged 85% over this period. The proportion of deaths under the age of 5 that were registered, decreased slightly from 77% to 75% over the period. It was thus decided, for national deaths, completeness remained constant at 85% for those under the age of 1 and at 78% for those under the age of 5 for 2006 to 2010.

As for the national estimates, completeness of reporting deaths by province, in the years 1997, 2001, 2006 and 2010 were calculated to reproduce the estimates of q_0 and ${}_5q_0$ in those years. (Estimates of q_0 and ${}_5q_0$ for 1997 and 2010 were linearly interpolated from the estimates spanning those years, namely, 1996 and 2001 for 1997 and 2006 and 2011 for 2010.) In addition, as for the national estimates, completeness for 1998 was set as the completeness for 1997 scaled up by the ratio of the number of vital registration deaths under the age of 1 in 1998 to those in 1997, because in many provinces there was a sharp increase in numbers that was unlikely to be due to an increase in mortality rates.

Estimates of completeness for the intervening years were derived by linear interpolation between the values for 1998, 2001, 2006 and 2010. These values and the implied series of estimates of q_0 and ${}_5q_0$ were then inspected for reasonableness and adjusted where necessary. In particular, a sudden increase or decrease from one year to the next in the reported deaths for no particular reason (as occurred from 1997 to 1998 in some provinces) was interpreted as a change in completeness rather than a change in rates. A general pattern of roughly level completeness from 1998 to 2001 followed by rapidly increasing completeness to 2006 and then declining mortality since 2006 was considered a feature of all provinces. In addition, the estimates of the rates for three provinces (Eastern Cape, Gauteng and Mpumalanga) appeared

to peak too sharply and so the linear trend in completeness between 2001 and 2006 was adjusted marginally to flatten out the peak.

In addition to the above, specific adjustments to trend in completeness were made to allow for:

i. Western Cape: A sharp increase in deaths from 1997 to 1999.
ii. Eastern Cape: A sharp fall in deaths from 2008 to 2009 and an increase from 2009 to 2010.
iii. Northern Cape: A sharp increase in deaths in 2007 and 2008.
iv. Free State: Excess deaths in 2008.
v. North West: A significant decline after 2006.

The provincial estimates of completeness in each year were adjusted proportionally to ensure that the sum of the estimates of deaths equalled the number of deaths for the country as a whole, after adjusting for completeness (on the assumption that national estimates of q_0 and ${}_5q_0$ are likely to be more reliable than those for individual provinces). Completeness factors for infant deaths are shown in Table A2 and for children aged from 1-4 years in Table A3.

Table A2: Percentage completeness of infant deaths by province and national, 1997-2010

	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
WC	101.6	109.3	115.6	115.0	114.2	105.8	101.1	97.6	95.4	91.9	97.9	103.1	115.2	121.5
EC	28.5	27.8	29.0	30.2	31.5	35.4	39.7	42.7	43.1	42.9	43.8	46.5	39.2	48.1
NC	85.9	93.0	90.8	89.1	87.4	83.7	82.6	82.5	83.4	83.4	109.4	111.6	120.9	123.8
FS	79.5	95.5	93.3	91.7	90.0	95.4	103.3	112.2	122.6	131.5	129.1	136.3	130.7	124.5
KZN	49.8	54.5	52.2	50.1	48.0	50.3	54.0	58.2	63.1	67.4	65.6	63.0	64.1	61.4
NW	68.8	84.6	85.5	86.8	88.0	95.4	105.2	116.0	128.3	139.1	131.8	133.8	119.8	110.2
GT	107.2	122.4	118.2	114.8	111.3	118.0	122.8	124.4	130.5	137.7	133.3	127.2	128.4	121.9
MP	40.0	51.8	49.6	47.8	45.9	54.1	59.1	64.4	70.6	76.0	74.3	71.7	73.4	70.7
LP	23.9	33.6	33.4	33.5	33.5	40.1	47.5	55.3	63.8	71.6	74.4	76.5	83.6	86.3
SA	54.5	62.2	62.0	61.8	61.6	66.3	71.0	75.6	80.3	85.0	85.0	85.0	85.0	85.0

Table A3: Percentage completeness of deaths of children aged from 1-4 at last birthday by province and national, 1997-2010

	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
WC	59.1	60.9	56.8	57.9	55.8	52.7	51.3	50.5	47.0	45.1	46.5	50.4	49.9	53.5
EC	22.0	26.0	27.9	30.0	31.9	31.3	33.5	35.8	37.0	37.6	41.0	46.1	42.4	50.8
NC	65.4	68.4	67.7	66.4	64.6	64.8	64.9	67.9	70.2	70.9	111.3	106.8	107.3	101.4
FS	53.3	65.0	62.2	58.5	56.6	60.0	65.1	72.4	76.4	84.4	81.2	82.3	89.6	83.9
KZN	37.7	47.3	44.3	42.0	40.1	40.8	42.4	47.1	49.0	49.7	48.2	45.8	46.0	44.8
NW	58.7	65.7	62.7	61.3	58.5	65.9	74.7	85.4	93.1	105.4	94.2	89.0	82.0	80.9
GT	59.5	75.9	69.7	63.8	60.0	70.8	79.2	86.7	81.1	75.3	71.4	68.4	67.5	66.4
MP	30.3	36.3	37.7	39.6	41.7	43.4	49.2	53.9	58.4	62.6	63.0	63.1	67.1	66.8
LP	35.4	45.9	47.5	46.7	47.4	54.1	59.9	67.2	77.2	84.8	90.6	93.6	101.7	99.1
SA	40.2	48.8	47.9	47.0	46.2	49.1	52.6	57.1	59.1	62.6	62.0	62.7	62.3	64.1

2.4 Determining the completeness of adult deaths by year

The completeness of adult death registration was derived using death-distribution methods and fitting logistic curves (one for each sex) to estimates of completeness (relative to the census/survey populations) for each year between the 1996 and 2001 censuses as produced by Dorrington, Moultrie and Timæus (2004), and estimates of completeness derived for the period between the 2001 census and the 2007 Community Survey. Annual estimates of completeness for the period 2001-2007 were derived on the assumption that completeness changed linearly over the period starting from the completeness estimated by Dorrington, Moultrie and Timæus for 2001. The logistic curve was fitted to these data with an asymptote of 95% completeness. As the curves for each sex crossed over and it was difficult to make a case for separate levels of completeness of reporting deaths of men and women, it was decided to blend the curves using blending weights that progressed linearly with time, such that completeness reached 93% by 2009. The completeness for ages 15+ was set at 93% for years beyond 2009.

Although somewhat arbitrary, the figure of 93% was chosen as being likely to be close enough to the true completeness from 2009 onwards. To check this assumption, the completeness was estimated for the period between the 2001 and 2011 censuses using the census populations, VR deaths for 2001 to 2010 and estimated VR deaths for 2011 from deaths recorded on the Population Register as part of the Rapid Mortality Surveillance project

(Dorrington, Bradshaw and Laubscher, 2014). The estimate of average completeness for this period was around 94%.

2.5 Determining the completeness of adult deaths by province

As application of the death-distribution methods to sub-national populations is unreliable, a different approach was needed to estimate the completeness of reporting adult deaths at a provincial level. First, unadjusted mortality rates were calculated for men and women aged 15+ for each province for each year from 1997 to 2010 by dividing the registered deaths for each province by an estimate of the population aged 15+ derived from the alternative mid-year estimates (Dorrington, 2013). Rates were also estimated for 2011 by apportioning the estimate of registered deaths of the population aged 15+ nationally, derived from the Rapid Mortality Surveillance data (Dorrington, Bradshaw and Laubscher, 2013), to the provinces in proportion to the numbers of deaths recorded in 2010.

Second, mortality rates were calculated for men and women aged 15+ for each province for the 12 months prior to the 2001 and 2011 Censuses and the Community Survey, using the deaths reported by households corrected for under/over-reporting, and the population aged 15+ derived from the alternative mid-year estimates (Dorrington, 2013). Estimates of under- or over-reporting deaths by households in the census was estimated by employing the same method as that used by Dorrington, Moultrie and Timæus (2004), namely to derive correction factors at the national level as the ratio of the numbers by sex, population group and age group from vital registration corrected for under-registration to the numbers reported by households in the census, on the assumption that all under-registration in the vital registration was confined to the rural, and hence the African, population. These estimates of the correction factors at the national level (by sex, population group and age group) were then used to adjust the number of deaths reported by households (by age, sex and population group) in each province for under- (or over-reporting).

As the trend in rates estimated from the deaths reported by households appeared to be implausible for many provinces, it was decided to use these estimates merely to indicate level and to use the estimates from the ASSA2008 provincial models to indicate the trend in rates over time. Thus rates were estimated as being those from the ASSA2008 provincial model scaled up or down such that the sum of squared differences of the estimates from the estimates derived from deaths reported by household for 1996, 2001, 2006 and 2011, were minimised. (For this purpose, the estimates from each census were assumed to apply to the

calendar year of the census, and those for the Community Survey to the calendar year 2006.) Since there was no estimate for 1996, this was set equal to the estimate from the ASSA model.

The ratio of the rates estimated from the VR deaths and those for each year produced from scaling the ASSA2008 rates to fit the estimates from the deaths reported by households provides an estimate of the completeness of reporting of deaths by province and year.

Final estimates of completeness were derived by averaging the estimates for men and women and then rebalancing them by province so that the sum of the estimates of the completely reported deaths by province for each year was equal to the estimate of the number of completely reported deaths for the country as a whole. These are presented in Table A4.

Table A4: Percentage of completeness of adult deaths (15+ years) by province and national, 1997-2010

	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
WC	109.7	113.6	113.9	108.4	108.2	106.8	101.8	100.2	101.9	97.1	101.5	102.5	99.0	100.4
EC	57.7	62.1	63.8	70.3	73.8	78.1	81.1	87.1	92.1	90.7	85.6	85.0	86.1	87.4
NC	78.5	83.1	78.1	78.8	78.5	81.1	81.1	83.5	80.7	82.1	100.7	103.9	105.5	111.4
FS	88.6	96.3	93.7	94.3	95.1	96.1	98.3	101.7	100.5	101.7	102.2	101.3	103.6	107.7
KZN	72.3	77.9	76.4	79.0	80.2	81.2	81.3	80.2	81.5	82.1	83.8	85.1	85.2	83.4
NW	81.9	89.8	93.7	96.8	97.1	100.2	104.2	105.5	107.4	108.8	96.2	96.0	87.7	92.1
GT	124.3	124.6	112.9	109.0	105.3	102.5	99.3	97.6	93.7	94.0	97.7	96.5	97.7	95.0
MP	69.9	73.4	73.9	74.0	76.0	75.7	79.2	78.9	79.0	81.2	86.4	86.2	87.9	88.6
LP	57.1	69.3	69.6	73.1	80.1	84.9	94.0	97.4	100.6	110.1	104.9	108.9	111.1	107.3
SA	80.2	85.8	84.3	86.0	87.4	88.7	89.8	90.7	91.4	92.0	92.5	92.9	93.0	93.0

THE SOUTH AFRICAN MEDICAL RESEARCH COUNCIL
Second National Burden of Disease Study for South Africa