CO-AG NUMBER: NU2GGH001150
CO-AG NAME: SOUTH AFRICAN MEDICAL RESEARCH COUNCIL

SOUTH AFRICAN

NATIONAL CAUSE-OF-DEATH VALIDATION PROJECT

REPORT 1 | METHODOLOGY AND DESCRIPTION OF A NATIONAL SAMPLE OF VERBAL AUTOPSIES

SAMRC Burden of Disease Research Unit | July 2020

PEPFAR
CDC Centers for Disease Control and Prevention
saMRC advancinglife
GeoSpace measure it. manage it.
Bloomberg Philanthropies
DATA FOR HEALTH INITIATIVE

South African National Cause-of-Death Validation Project:
Methodology and Description of National Sample of Verbal Autopsies

SUGGESTED CITATION

Bradshaw D, Joubert JD, Maqungo M, Nannan N, Funani N, Laubscher R, Cheyip M, Zinyakatira N, Awotiwon O, Nojilana B, Bezuidenhout F, Martin L, Dempers J, Kahn K, Price J, Lombard C, Morof D, Nichols E, Rao C, Groenewald P. South African National Cause-of-Death Validation Project: Methodology and Description of a National Sample of Verbal Autopsies. Cape Town: South African Medical Research Council, 2020.

COLLABORATORING ORGANISATIONS

South African Medical Research Council
Statistics South Africa
National Department of Health
Department of Home Affairs
Centers for Disease Control and Prevention
South African Funeral Practitioners Association
National Funeral Directors Association

PRINCIPAL INVESTIGATORS

Prof Debbie Bradshaw, Dr Pam Groenewald, Dr Jané Joubert

CO-INVESTIGATORS

Prof Chalapati Rao, Ms Cherie Cawood, Dr Estevão Afonso, Mr Francois Bezuidenhout, Dr Jessica Price, Prof Johan Dempers, Dr Kassahun Ayalew, Prof Kathleen Kahn, Prof Lorna Martin, Dr Megan Prinsloo, Ms Mireille Cheyip, Dr Nadine Nannan, Mr Nesbert Zinyakatira, Prof Sam Clark, Dr Tshilidzi Muthivhi, Mr Thabo Molebatsi

TECHNICAL ADVISORS

Dr Beatrice Nojilana, Prof Carl Lombard, Dr Diane Morof, Dr Erin Nichols, Dr Frank Odhiambo, Dr Ian Neethling, Dr Lyn Hanmer, Ms Mmamokete Mogoswane, Ms Mosidi Nhlapo, Prof Sam Notzon, Ms Sizzy Ngobeni, Dr Oluwatoyin Awotiwon, Dr Victoria Pillay-van Wyk

PROJECT COORDINATORS

Dr Monique Maqungo and Ms Noluntu Funani

IMPLEMENTATION PARTNERS

Phase 1: Epicentre Health Research
Phase 2: GeoSpace International

FUNDING

This study has been supported in part by the President's Emergency Plan for AIDS Relief (PEPFAR) through the Centers for Disease Control and Prevention (CDC) under the terms of Cooperative Agreement SAMRC-CDC CoAG 1U2GGH01150. It has also been supported by the Bloomberg Philanthropies Data for Health Initiative through the CDC Foundation Civil Registration and Vital Statistics Project, and the South African Medical Research Council.

DISCLAIMER

The content of the report is that of the authors and does not necessarily represent the official position of the funding agencies.

ACKNOWLEDGEMENTS

The fieldwork teams from Geospace and Epicentre are thanked for the data collection conducted across South Africa.
Professor Peter Byass is thanked for his technical advice on InterVA for automated processing of verbal autopsy data and Ms Natashia Morris of the SAMRC Biostatistics Unit is thanked for the maps that she created for the project.

A copy of this resource is available at: www.mrc.ac.za/bod/reports.htm

CONTENTS

List of Tables

List of Figures

Abbreviations and Acronyms

AIDS	Acquired immune deficiency syndrome
CDC	Centers of Disease Control and Prevention
CI	Confidence interval
COD	Cause of death
COMCATs	Circumstances of mortality categories
CRVS	Civil Registration and Vital Statistics
CSMF	Cause specific mortality fraction
DHA	Department of Home Affairs
DNF	Death notification form
FP	Funeral Practitioner
FPS	Forensic Pathology Services
HDSS	Health and demographic surveillance system
HIV	Human immunodeficiency virus
ICD-10	International Classification of Diseases Tenth Revision
ICS	International Classification of Diseases and Related Health Problems
MIA	Minimally invasive autopsy
NBD	National burden of disease
NCODVP	National cause of death validation project
NOK	Next of kin
ODK	OpenDataKit
QA	Quality assurance
RTHC	Road to health card
SA ID	South African Identification Number
SA NBD	South African National Burden of Disease Study
SAMRC	South African Medical Research Council
SDG	Sustainable Development Goals
SES	Socio-economic status
SOP	Standard operating procedure
Stats SA	Statistics South Africa
TB	Tuberculosis
UCOD	Underlying cause of death
USID	Unique study identification
VA	Verbal autopsy
WHO	World Health Organisation

GLOSSARY

Aggregation of causes of death

The analysis of the causes of death in this report makes use of the 10th revision of the International Statistical Classification of Diseases and Related Health Problems (ICD-10). This is a standardized medical classification list by the World Health Organization (WHO), updated in 2016. It classifies diseases and related health problems into 22 chapters, of which 19 are used in the reporting of information on underlying causes of death. (Available at https://icd.who.int/browse10/2016/en).

A number of lists of aggregated causes have been developed for working with verbal autopsy data. This report uses the 2016 cause of death list for verbal autopsy comprising 64 causes mapped onto ICD-10. (Available at https://www.who.int/healthinfo/statistics/verbalautopsystandards/en/).

Further analysis has been done by grouping the ICD-10 causes and the VA cause list into 3 broad cause groups with an additional category for HIV/AIDS and TB as has been used in the South Africa Burden of Disease studies. These are:

- HIV/AIDS and TB
- Other infections
- Non-communicable diseases
- Injuries.

Cause of death sequence

The cause of death sequence is the chain of events leading directly from the underlying cause to the immediate cause of death.

Circumstances of mortality categories (COMCAT)

Developed by Hussain-Alkhateeb *et al* (2019), these categories describe the circumstantial determinants of death which can be assessed in parallel with medical cause. The categories are automatically identified by InterVA-5 based on the questions from the WHO2016 Verbal Autopsy:

- *Traditions* – Traditional practices or beliefs influenced health-seeking behaviour and the pathway to death
- *Emergencies* – Sudden, urgent or unexpected conditions leading to death, which probably precluded life-saving actions
- *Recognition* – Lack of recognition or awareness of seriousness of disease (e.g. symptoms or severity) negatively influenced health seeking behaviour
- *Resources* – Inability to mobilise and use resources (e.g. material, transport, financial) hindered access to care
- *Health Systems* – Problems in getting health care despite accessing health facilities (e.g. related to admissions, treatment and medications)
- *Inevitability* – Death occurred in circumstances that could not reasonably have been averted (e.g. very elderly or recognised terminal conditions)
- *Multiple* – A combination of the above categories affected the pathway to death: no single factor predominated.

Community Oriented Primary Care

Community oriented primary care (COPC) is a strategy whereby elements of primary health care and of community medicine are systematically developed and brought together in a coordinated practice.

Death

The permanent disappearance of all evidence of life at any time after a live birth has taken place, or postnatal cessation of vital functions without capability of resuscitation. This definition excludes fetal deaths, i.e. stillbirths (see definition below).

Decedent/deceased

Persons who died in South Africa and whose body has been taken to a designated funeral parlor registered with the Department of Home Affairs, or whose body has been prepared for burial or cremation by a funeral undertaker, or whose death has been registered directly at a local Department of Home Affairs office by a next of kin/carer/friend of the decedent. Foreigners who died in the country were included in the study when an adult (18 years+) next of kin/carer/friend could be contacted within the study timeframe and could speak English or any of the nine most common South African official languages into which verbal autopsy questions were translated.

ICD-10

The International Classification of Diseases and Related Health Problems (ICD) is a classification and coding system developed by the World Health Organisation (WHO) and defines the universe of diseases, disorders, injuries and other related health conditions, listed in a comprehensive, hierarchical fashion. The 10th revision, updated in 2016, is currently used as the international standard for reporting diseases and health conditions and can be found online. The next revision of ICD has been completed and it is anticipated that over the next few years, ICD-11 will be adopted.

Injury death

Deaths due to injuries (external causes) are required by law in South Africa to undergo a post mortem investigation at Forensic Pathology Services to determine culpability and cause of death.

International Form of Medical Certificate of Cause of Death

The ICD has outlined principles for certifying the medical cause of death and the rules for coding which are essential for standardising cause of death statistics. This starts with the form that has a specific layout and needs to be completed in a specific way to ensure that the underlying cause of death can be identified.

The sequence of the causes of death from the underlying cause to the immediate cause should be reported in part I of the form with immediate cause shown in line (a). Other conditions that contributed to the death should be reported in part II.

	Cause of death	Approximate interval between onset and death
I		
Disease or condition directly leading to death*	a) ..... due to (or as consequence of)	.....
Antecedent causes		
Morbid conditions, if any, giving rise to the above causes, stating the underlying condition last	(b) ..... due to (or as consequence of)	.....
	(c) ..... due to (or as consequence of)	.....
	(d)	.....
II		
Other significant conditions contributing to the death, but not related to the disease or condition causing it	.....	.....
	.....	.....
** This does not mean the mode of dying, e.g., heart failure, respiratory failure. It means the disease, injury, or complication that caused death.*		

The sequence of the causes of death from the underlying cause to the immediate cause should be reported in part I of the form with immediate cause shown in line (a). Other conditions that contributed to the death should be reported in part II.

Iris

Iris is an automated system for coding multiple causes of death and for the selection of the underlying cause of death based on the ICD-10 coding rules. It can be used in batch or interactively.

InterVA

InterVA is a suite of computer models to facilitate interpreting verbal autopsies towards generating a probable cause of death, using a Bayesian approach. The latest version InterVA-5 has been used in this project.

Medical doctor/physician

A medical doctor is a trained health professional who practices medicine, which is concerned with promoting, maintaining, or restoring health through the study, diagnosis, prognosis and treatment of disease, injury, and other physical and mental impairments. The term 'medical doctor' is used interchangeably with 'physician' in this report.

Multiple causes of death

When coding and classifying causes of death, you must first assign ICD codes to all the conditions reported on the death certificate. Many coding instructions are based on specific ICD codes and, to determine whether any of the instructions apply, you need to know the ICD codes for all conditions on the certificate. This is called multiple-cause coding.

Next of Kin (NOK)

The deceased's close living relatives are known as the next of kin and in this report, the informant is the person who reported the death to the Department of Home Affairs.

Ninety-five percent confidence interval (95% CI)

The 95% confidence interval represents the sampling variability around an estimate. A 95% confidence interval (CI) of a statistic is a range with an upper and lower number calculated from a sample that describes possible values that the true statistic could be. If multiple samples were drawn from the same population and a 95% CI calculated for each sample, we would expect the population statistic to be found within 95% of these CIs.

Stillbirths

The definition recommended by WHO for international comparison is a baby born with no signs of life at or after 28 weeks' gestation. A **fresh stillbirth** is defined as the intrauterine death of a fetus during labor or delivery, and a **macerated stillbirth** is defined as the intrauterine death of a fetus sometime before the onset of labor, where the fetus showed degenerative changes.

Sustainable Development Goals (SDG)

The Sustainable Development Goals, also known as the Global Goals, were adopted by all United Nations Member States in 2015 as a universal call to action to end poverty, protect the planet and ensure that all people enjoy peace and prosperity by 2030. Cause of death data are a prerequisite to measure several indicators.

Underlying cause of death (UCOD)

The underlying cause of death, from a public-health point of view, is considered the most informative cause-of-death-data element, and therefore was designated the cause of death for primary tabulation and comparisons. From the perspective of prevention of death, "it is necessary to break the chain of events or to effect a cure at some point. The most effective public health objective is to prevent the precipitating cause from operating. For this purpose, the underlying cause has been defined as "(a) the disease or injury which initiated the train of morbid events leading directly to death, or (b) the circumstances of the accident or violence which produced the fatal injury".[1] To properly select the underlying cause of death, coders are taught to apply the *International Statistical Classification of Diseases and Related Health* (ICD) rules and instructions to the sequence of causes as indicated on the *International Form of Medical Certificate of Cause of Death*. Automated software developed by the Iris Institute is available to facilitate coding of multiple causes of death and selection of the correct underlying cause.

Unusable code

Unusable codes (also referred to as 'garbage codes') are any ICD code that cannot or should not be considered an underlying cause of death, such as septicaemia, senility or headache. They may also be the code for a cause that belongs in some other part of the morbid sequence of events leading to death such as the immediate or intermediate cause; or a cause of death that is insufficiently specified. Essentially, an unusable code is one that has no use in informing public health policy, as the related underlying cause of death (UCOD) is too vague, or simply impossible. Mikkelsen et al (2017) have defined five categories of unusable codes in the ANACONDA tool:

- Category 1 – Symptoms, signs and ill-defined conditions
- Category 2 – Impossible as underlying causes of death
- Category 3 – Intermediate causes of death
- Category 4 – Immediate causes of death
- Category 5 – Insufficiently specified causes within ICD chapters.

Verbal autopsy (VA)

A method of determining an individual's cause/s of death using a trained interviewer to administer a questionnaire during a face-to-face interview to collect information about the signs, symptoms, treatment, and demographic characteristics of a recently-deceased person from another individual – ideally a close caregiver or family-member – with knowledge about the deceased during his/her terminal illness/event.

Ward Based Outreach Teams (WBOTS)

A team of community health workers (10-20) with a team leader (professional or enrolled nurse) who are responsible for primary health care service delivery in a defined municipal ward comprised of about 200 households.

EXECUTIVE SUMMARY

South Africa National Cause of Death Validation Project

South Africa has a well-established Civil Registration and Vital Statistics System (CRVS) with a high proportion of deaths being registered. The quality of the cause of death statistics that arise from the death notification forms completed by medical doctors, however, is sub-optimal, making it difficult to obtain a reliable cause of death profile or monitor trends in the number of deaths from specific causes. In addition, there is extensive underreporting of HIV as an underlying cause of death. Although estimates based on demographic models of the epidemic suggest that 17% of the deaths in that 2016 were due to HIV, the country's vital statistics report indicated that only 4.8% of all the registered deaths were due to HIV.

The South Africa (SA) National Cause of Death Validation Project (NCODVP) was implemented by the South African Medical Research Council (SAMRC) and partners to conduct a validation of CRVS cause-of-death information by linking CRVS data to data obtained from medical records, forensic pathology records, and verbal autopsy interviews for a national sample of deaths. The main purpose of the study was to compare the registered underlying cause of death indicated on the CRVS medical notification form with the highest level of information collected in the study (forensic pathology record followed by medical records and verbal autopsy) so that correction factors could be estimated. Additionally, the study aimed to compare the medical cause of death identified from the different sources to assess their performance in identifying cause of death. The study protocol was reviewed and approved by the SAMRC Ethics Committee and the United States Centers for Disease Control and Prevention (CDC)Centers for Global Health Associate Director for Science. Support was obtained from the National Department of Health and permissions were obtained from each provincial Department of Health and health facility included in the study.

Purpose of this report

- This first project report outlines the rationale, aims and objectives of the study. The report also describes the study methodology and realisation of the study sample of deaths.
- Initial results from the verbal autopsies, including determination of underlying cause of death based on 1) InterVA-5 automatic selection from standardized verbal autopsy data and 2) determination based medical certification of the causal sequence of death after review of the verbal autopsy data by a doctor and coded to ICD-10 using Iris software. Both cause of death results are compared with the national cause of death data compiled by Statistics South Africa (Stats SA) for 2016.

Future reports

- A second report is planned once the medical records and forensic pathology records have been processed and a third report will be prepared once the sample data have been linked to the CRVS data.

Study design and method

A sample size of >13,000 deaths from 27 randomly selected sub-districts across the country was assessed to provide sufficient precision for the correction factors for deaths caused by four selected conditions including HIV, cerebrovascular disease, diabetes mellitus, and interpersonal violence.

- In the first phase of fieldwork that took place during 1 August 2017 to 30 April 2018, funeral undertakers and Department of Home Affairs offices were enrolled to inform decedents' next of kin about the study and seek permission for the research team to contact them at a later stage. A fixed 3-month census period (1 September 2017 to 30 November 2017) was planned to provide the target number of decedents for validation, but the study period needed to be extended to nearly 8 months (1st September 2017 to 13th April 2018) due to low recruitment. By the end of the first phase of fieldwork, a total of 6,328 next of kin had consented to be approached by the research team, accounting for 65% of the 9,731 who were approached.

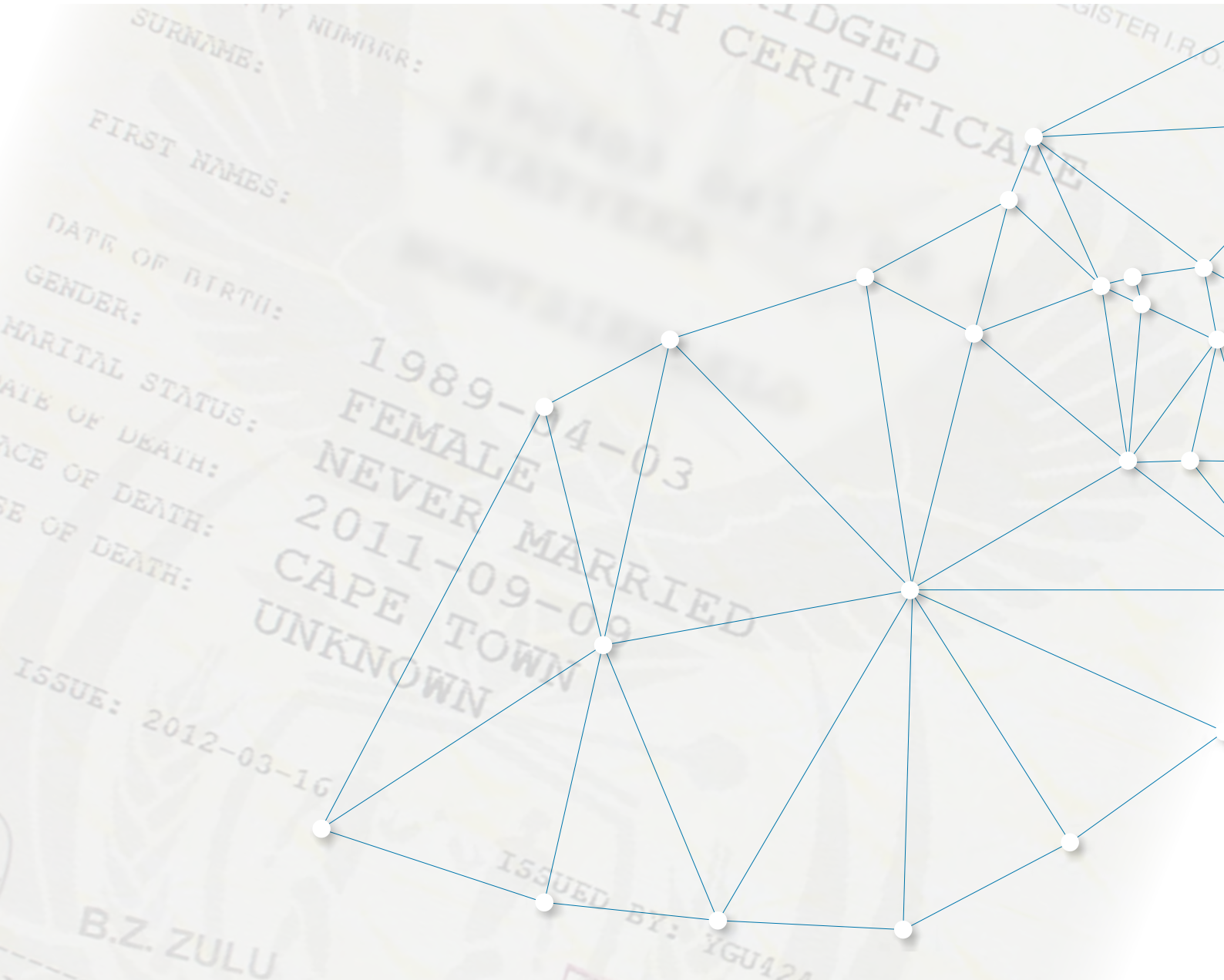
- In the second phase of fieldwork that occurred during August 2018 – March 2019, trained fieldworkers contacted the next of kin to arrange to conduct a face-to-face verbal autopsy interview using the three questionnaires of the WHO 2016 instrument[2] that had been translated into nine official languages. In addition, all medical records and forensic pathology records of decedents who passed away in the facilities within the selected sub-districts during the study census period (1 September2017 to 13 April 2018) were anonymised, scanned, and submitted to the office-based quality control team (GeoSpace International) for review. Thereafter, the anonymised records were submitted to the investigators.
- After the completion of fieldwork, data obtained from the Department of Home Affairs indicated that 36,976 deaths were registered with place of occurrence in the 27 sampled sub-districts during the study census period 1 September 2017 – 13 April 2018. A total of 5,387 verbal autopsies were successfully conducted, and 17,625 medical records and 5,742 forensic pathology records collected. Information was collected for 26,514 decedents yielding a ratio of 72% relative to the target population of registered deaths and well over the number of deaths identified in the sample size determination
- While the apparent overall response rate of deaths that included data from at least one source was high (72%), the response rate for verbal autopsy interviews was very low (15%). Furthermore, the place of occurrence of some of these deaths was outside the sampled areas. Of the 4,928 deaths for which we had the SA identification number, 58.9% could be matched to the target sample that was provided by the Department of Home Affairs some months later. Some of the sampled areas were particularly difficult to delineate as the Department of Health uses Municipal boundaries while the Department of Home Affairs uses magisterial district boundaries. Relative to the target population of all registered deaths within the study area, the realised sample of deaths in Gauteng and Eastern Cape provinces are under-represented while the other provinces are over-represented. The potential for bias in the realisation of the sample will need to be assessed once all the data are processed and linked with the CRVS data, and it is possible that a weighted analysis will need to be undertaken.

Verbal autopsy interview results

- Of the VAs conducted, 68.3% were assigned a specific and valid underlying cause of death, and 16.6% were assigned an underlying cause of death within an ICD chapter without sufficient specification (e.g., cancer without primary site). Although 16.6% of the causes were identified to have insufficient specification within an ICD chapter, this should be anticipated as the information was derived from a verbal autopsy tool based on lay-person reporting
- HIV was the most common underlying cause of death in this sample of verbal autopsies, accounting for 22.8% of all deaths. Out of the 1,601 deaths from HIV and TB, 10.2% had mention of HIV treatment default while 4.0% of the TB deaths had mention of treatment default. Using InterVA-5 to automatically identify the most probable cause of death also revealed HIV related causes as the most common.
- When deaths from this study were compared with the national cause of death data compiled by Statistics South Africa (Stats SA) for 2016, the latest year for which they are available, it was observed that the Stats SA data had a similar proportion of injury deaths, a much lower proportion of HIV and TB deaths, and higher proportions of unknown causes, non-communicable diseases, and other group1 conditions (communicable diseases: other infections, maternal, neonatal, and nutritional conditions).

Key findings and recommendations

- The main finding of the project thus far has been the demonstration that conducting good quality verbal autopsies on a national scale is feasible, however, the identification and recruitment of next of kin is complex and might be best achieved in a local setting or else be part of a regulated service through the Department of Health or Home Affairs. Furthermore, this system demonstrated that HIV can be identified as a cause of death through verbal autopsies.
- The processing of the medical records and forensic pathology records needs to be completed and the data linked with CRVS data to estimates correction factors, the main objective of the study.
- The data from the physician reviewed verbal autopsies should be analysed further to describe the cause-symptom patterns found in the South African setting and used to enhance the available the tools for automated selection of cause of death such as Inter-VA. It is important to explore the performance of the verbal autopsy instrument. In addition, these data need to be used by researchers together with verbal autopsy data collected in other countries to enhance the verbal autopsy questionnaires.
- The resources from the project that were developed to train doctors in medical certification need to be used in a national effort to improve the quality of medical certification.



1. INTRODUCTION

1.1 The importance of mortality data for health

Reliable, continuous, and timely mortality data, and valid cause-of-death information, by age and sex, are widely acknowledged as essential intelligence for national governments to detect, prioritize and address challenges that stand in the way of improving population health and enhancing people's survival.[3,4] National age-sex mortality rates and changes over time in the leading causes of death are critical information required to alert decision-makers to changes in cause-of-death patterns and trends that are required for adjusting priorities and distribution of public health resources.

The ideal source of a country's mortality data is a well-functioning, national, full-coverage civil registration and vital statistics (CRVS) system with high levels of completeness of death registration, thorough ascertainment of the cause/s of death by medical doctors well-trained in the medical certification of the cause of death, and timely-published vital statistics reports.[5-7] The statistical, administrative, and legal advantages of complete and accurate civil registration systems over other mortality-data collection systems such as sample surveys, censuses, sample registration systems, or disease registries,[8] have been acknowledged for decades by international agencies such as the United Nations,[6,7,9-14] World Health Organization (WHO),[15-19] World Bank,[20,21] and throughout the iterations of the Global Burden of Disease (GBD) studies, the Institute for Health Metrics and Evaluation (IHME).[4,22-24] More recently, the 2030 Agenda for Sustainable Development[25,26] clearly illustrates the importance and advantage of countries having a national CRVS system in that 67 of 230 proposed indicators to monitor progress in 12 of the 17 total Sustainable Development Goals (SDG) can be measured from data derived from well-functioning CRVS systems.

The prominence of mortality reduction among the health-related SDGs has intensified countries' need for robust national mortality measurements to monitor levels and causes of mortality.[4] Complete and accurate CRVS systems would be the best source of all-cause and cause-specific mortality data to measure progress in eight SDG targets related to reducing mortality, including maternal mortality ratio (indicator 3.1.1); all-cause under-5 (3.2.1) and neonatal (3.2.2) mortality; premature mortality from non-communicable disease (3.4.1); and mortality rates from suicide (3.4.2), homicide (16.1.1), road traffic injuries (3.6.1), and poisoning (3.9.3).[8,26]

At sub-national levels, robust and continuous regional or small-area patterns and trends of the levels and causes of death are equally important to identify and address health inequalities and differential access to health services. South African vital statistics, disaggregated at regional and smaller administrative levels, and carefully interpreted in accordance with expert medical and epidemiological input, have highlighted provincial,[27,28] health district,[29] and local-level[27,30,31] inequalities.

1.2 Civil registration and vital statistics in South Africa

Death registration was enacted in South Africa as long ago as 1867, and the national statistical office was established in 1914. However, ideologically-induced differential civil registration practices existed for different geographic areas, population groups, and residential statuses, contributing to partial coverage and low completeness of death registration for most of the 1900s.[32] South Africa has made great strides in increasing geographic coverage of death registration,[33] in particular after the geo-political integration of the "homeland" areas and the enactment of the Births and Deaths Registration Act of 1992,[34] which theoretically left no scope for optional or differential registration.[35] These events, and pivotal collaborative work by mortality scientists, Statistics South Africa (Stats SA), and the National Departments of Health and Home Affairs, described elsewhere,[27,33,36-47] facilitated rapidly-increasing levels of completeness among adults (~93%), as well as among infants (currently approximately 70%) and children aged 1-4 years (~65%).[35,48-51]

Currently, South Africa houses a well-functioning, inter-operable civil registration, vital statistics and identity management system, settled within a legal framework provided by the Births and Deaths Registration Act (Act no 51 of 1992).[34] The act requires all deaths and stillbirths to be notified to the Department of Home Affairs (DHA) on the official death notification form, Form DHA-1663, which includes the 2010 revision of the WHO international form of the Medical Certificate of Cause of Death. As indicated in Figure 1, a medical doctor or forensic pathologist is required to complete and sign the form. All deaths arising from unnatural causes must be referred to forensic pathology services for certification. In areas with poor access to physicians, notice of a death may also be given via the Death Report (DHA-1680) provided that the death was due to a natural cause. The latter may be filled by authorized traditional leaders/headmen, members of the South African Police Service, and undertakers designated by the DHA.

From Figure 1, it can be seen that once a death notification form (DHA-1663) has been completed, a funeral undertaker may register the death at the nearest local DHA office, or the informant (usually kin, carer, neighbor, or friend of the decedent) will report the death directly to the local DHA office. The information is entered onto the electronic National Population Register and a burial order and abridged death certificate are issued. From here, the forms are sent to the DHA head office in Pretoria where the documents are verified. The death notification forms are then submitted to Stats SA, where the immediate, antecedent, and contributory causes of death are coded, analysed, and tabulated. An underlying cause of death (UC) gets assigned using the Iris software[52] and the guidelines of the WHO's tenth revision of the International Statistical Classification of Diseases and Related Health Problems (ICD-10).[53,54] Annual cause of death reports are released and anonymised unit record data are made available for further epidemiological analysis (http://www.statssa.gov.za).

1.3 Cause-of-death data quality challenges

Despite these improvements in death registration, major challenges remain with the way that doctors complete the medical certificate of the cause/s of death and the consequent quality of cause-of-death information. These include a high proportion of deaths with ill-defined causes (13%), and an additional 13% having a cause of death not valid as an underlying cause in 2016,[55] under-reporting and misclassification of HIV deaths and an inaccurate profile of injury deaths[56] (for example accidental gun deaths are too high and homicides are too low). The extent of these problems varies at district and sub-district levels.[57]

A number of small-sample urban validation studies have highlighted the inaccuracy of cause-of-death attribution in CRVS data, particularly regarding high proportions of HIV deaths being misattributed to other cause.[47,58-61] Misclassification of ischemic heart disease, diabetes, and hypertensive disease has also been found.[61] Both these patterns were also found in a rural study linking and comparing CRVS cause-of-death data and verbal autopsy diagnoses as reference diagnoses from the Agincourt Health and Socio-demographic Surveillance System (Agincourt HDSS), a surveillance site established in 1992 comprised of 31 villages and 120 000 people in rural northeast South Africa.[62] At the national level, modeling studies using CRVS mortality data from 1996 to 2006 confirmed large proportions of HIV deaths misattributed to pseudonyms for HIV such as "immune suppression" and other immediate causes of death without indication of HIV.[63,64]

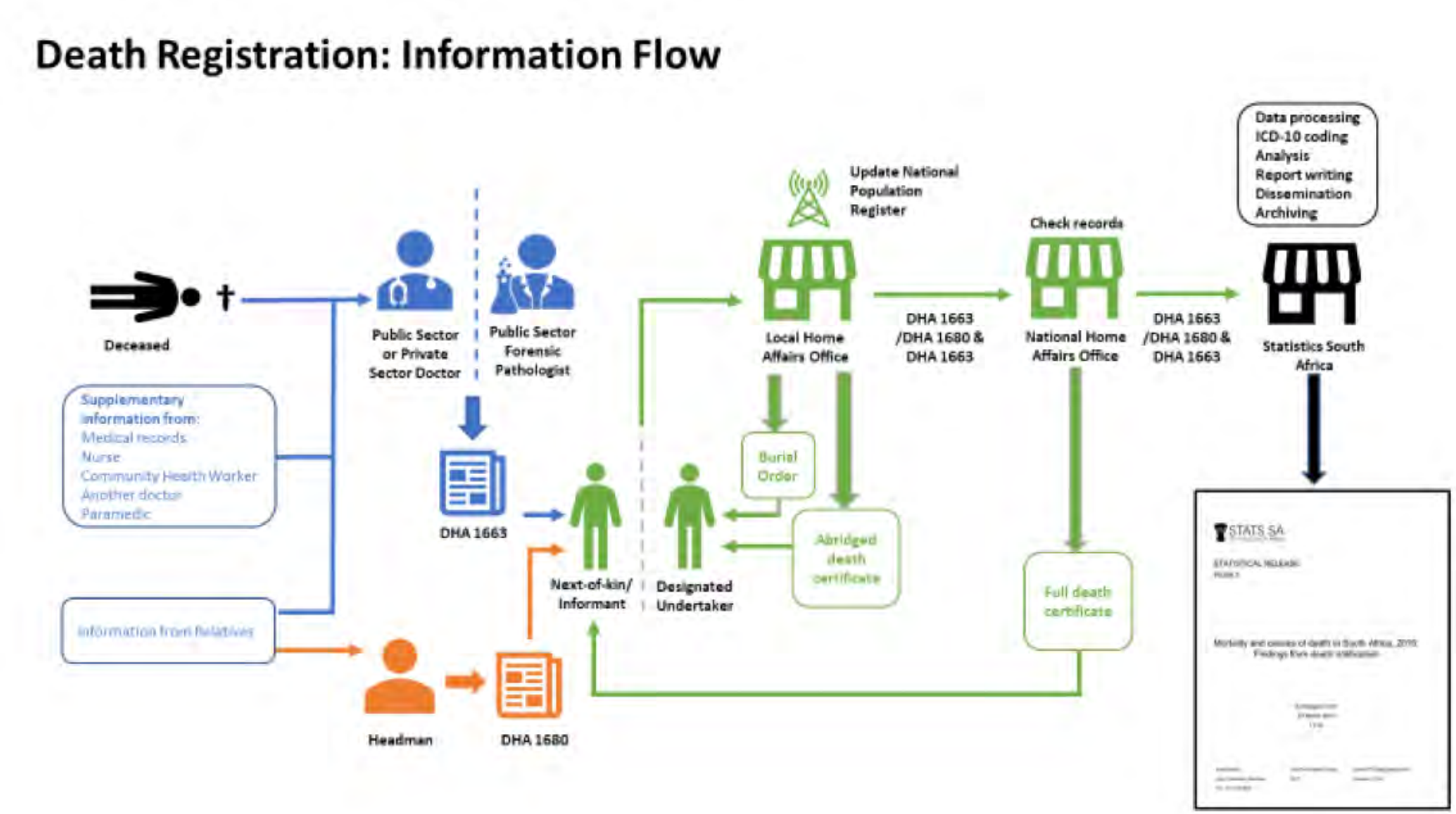


Figure 1: Schematic of information flow for death registration in South Africa.

In addition, a nationally representative survey of injury mortality investigated at mortuaries showed substantial misclassification of the manner and external causes of injury-related deaths in official mortality data,[65] which confirmed findings from the above-mentioned linkage study.[62]

1.4 Reference standards available in South Africa

Validating the causes of death on individual medical certificates from CRVS records requires a cause-of-death data source that can serve as a reference standard.[62,66,67] Post-mortem physical autopsies are generally considered the diagnostic gold standard against which the accuracy of reported causes of death can be ascertained.[68-70] However, these are expensive and can be culturally or religiously unacceptable to relatives and communities due to the invasiveness of the procedure.[71-74] As an alternative to a complete diagnostic forensic autopsy, the minimally invasive autopsy (MIA) was proposed in 2009.[75,76] In 2015, the simplified MIA, involving blood, cerebrospinal fluid, and organ-directed sampling using biopsy needles, was implemented by trained technicians in low-income settings, and found to be a successful procedure towards improved attribution of causes of death in developing countries.[69,77] In the absence of a full autopsy or MIA, a medical record review, where available, has been shown to provide a reasonable alternative for validating reported causes of death.[78,79]

Over the past 15 years, between 41% and 48% of annual deaths in South Africa occurred in health facilities[80] where there is an expectation that medical records would be available for the decedent. With more than half of annual deaths occurring outside health facilities, reference sources other than hospital record reviews are required for validation purposes. For injury deaths in South Africa, forensic autopsy records have been shown to provide a suitable reference source for attributing or validating causes of death.[65] For deaths that occur outside health facilities, study results from the Agincourt HDSS have illustrated that verbal autopsies can result in reliable cause-of-death results, despite acknowledged limitations, and that there is potential for verbal autopsy diagnoses to be used as a reference diagnosis for CVRS data.[62,81,85,86]

1.5 Rationale for a national cause-of-death validation project

Substantial misclassification of CRVS cause-of-death data have been documented, particularly for HIV, tuberculosis, injuries, and cardiovascular causes, as well as a large proportion of deaths certified with ill-defined/non-specific causes. Moreover, valid cause-of-death data are critical to inform health planning and prioritize and evaluate interventions aiming to improve population health and reduce health inequalities. Despite this knowledge, the validity of national CRVS cause-of-death data has not been studied in a nationally representative sample of deaths in South Africa.

A national validation study of cause-of-death statistics is critically important because deaths due to HIV/ acquired immunodeficiency syndrome (AIDS) and TB need to be accurately quantified, as these have become endemic[45,82] and were major contributors to the rapidly-reduced life expectancy seen until 2006,[22,36,64,83] and there are alternative mortality data sources that can be used to assess causes of death. These include hospital and forensic pathology records for facility and injury deaths, respectively, and the standardised WHO instruments for conducting verbal autopsies for deaths occurring outside health facilities.

2. AIMS AND OBJECTIVES

2.1 Aim

The overall aim of the NCOD validation project is to derive estimates of cause-specific mortality patterns in South Africa in 2017 at national, provincial, and district levels, using civil registration data validated and corrected against cause-of-death data from hospital, forensic, and verbal autopsy records.

2.2 Objectives

The study has three interrelated objectives with detailed sub-objectives described in Annexure 8.1.

The broad objectives of the project are:

1. To verify causes of death reported on CRVS death notification forms in a nationally representative sample of deaths occurring within and outside health facilities.
2. To derive correction factors to adjust cause-specific mortality data from CRVS according to reference diagnoses at national, provincial, and district levels.
3. To design and test a standardized methodology for household verbal autopsy for deaths occurring outside health facilities, with a view towards broader implementation within the routine CRVS system.

2.3 Purpose of report

Data collection has been completed and data processing is underway. This first project report outlines the study methodology and describes the sample realisation. This report also presents and discusses initial results from the verbal autopsies, including determination of underlying cause of death as determined by 1) InterVA-5 automatic determination based on standardized verbal autopsy data and 2) Doctor review of verbal autopsy questionnaire and narrative to determine cause of death sequence and causes coded using Iris software. Both cause of death results are compared with the national cause of death data compiled by Statistics South Africa (Stats SA) for 2016 as the 2017 data were not available at the time of analysis. A second report will provide the cause of death profiles from the medical and forensic pathology records and the third report will provide the results from the data linkage with CRVS data.

3. METHODS

3.1 Study design

This was a cross-sectional study using data collected for a fixed-period census of deaths that occurred in a nationally representative sample of health sub-districts in South Africa during part of 2017 and 2018 (see below). Verbal autopsy interviews were conducted through face-to-face interviews in the household of the next of kin/carer and medical records (MR) and forensic pathology service records (FPS) were collected from facilities serving the selected areas. Data were reviewed by trained doctors to identify the underlying cause of death. The underlying cause of death reported in the CRVS will be validated against the underlying cause identified through the highest level of evidence collected in the study for each decedent. The forensic pathology information will be considered the highest level of evidence, followed by the medical record and then the verbal autopsy.

3.2 Target population

The study population was the registered deaths in South Africa for the period 1 September 2017 until 13 April 2018. A limited period was selected to facilitate data collection. While a seasonal variation in disease profile is expected, there is no reason to anticipate that bias or distortions in the certification of cause of death would have a seasonal variation. To ensure that no deaths were missed, field collection included all deaths including stillbirths. The stillbirths have been described but not included in the report as deaths.

3.3 Sampling

A nationally representative random sample of 27 sub-districts (Figure 2) was selected using pseudo stratification according to socio-economic status (SES) within each province. The country is divided into 52 health districts, comprised of eight metropolitan municipalities and 44 municipal districts, across the nine provinces. Each district is divided into several sub-districts with a total of 234 across the country, including the 8 metropolitan municipalities. The sub-districts represent the lowest unit of health management within the tiered administration of the public health system in South Africa. The sampling frame for the NCOD validation project in South Africa was based upon a modification of 234 sub-districts and comprised of 226 municipal sub-districts and 2 metropolitan municipalities (Buffalo City and Mangaung). As the six metropolitan municipalities of Cape Town, Nelson Mandela Bay, Johannesburg, Ekurhuleni, Tshwane and eThekweni are populous and have varied socio-economic conditions, they were sub-divided into their 34 health sub-districts, increasing the sampling frame to 262. For practical reasons, sub-districts with less than 100 deaths per annum were pooled with a contiguous sub-district. These included combining Laingsburg and Prince Albert; Ikwezi and Baviaans; Kamiesberg and Khai-Mai; and Mier and Khai Garieb, resulting in a final sampling frame of 258 sub-districts (Annexure 8.2).

The pseudo-stratification according to SES was achieved by ranking the health sub-districts by the poverty headcount as assessed in Census 2011[84] and dividing them into tertiles based on SES rank (Annexure 8.2). One health sub-district (primary sampling unit) was then randomly selected from each tertile of the sub-districts within a province. Within each selected sub-district, a fixed period census of deaths was planned to attain an anticipated sample of 13,000 deaths, based on the overall number of expected annual deaths in the 27 sampled sub-districts. The period required to provide the required sample size was determined based on the expected numbers of deaths in the selected sub-districts (Annexure 8.3).

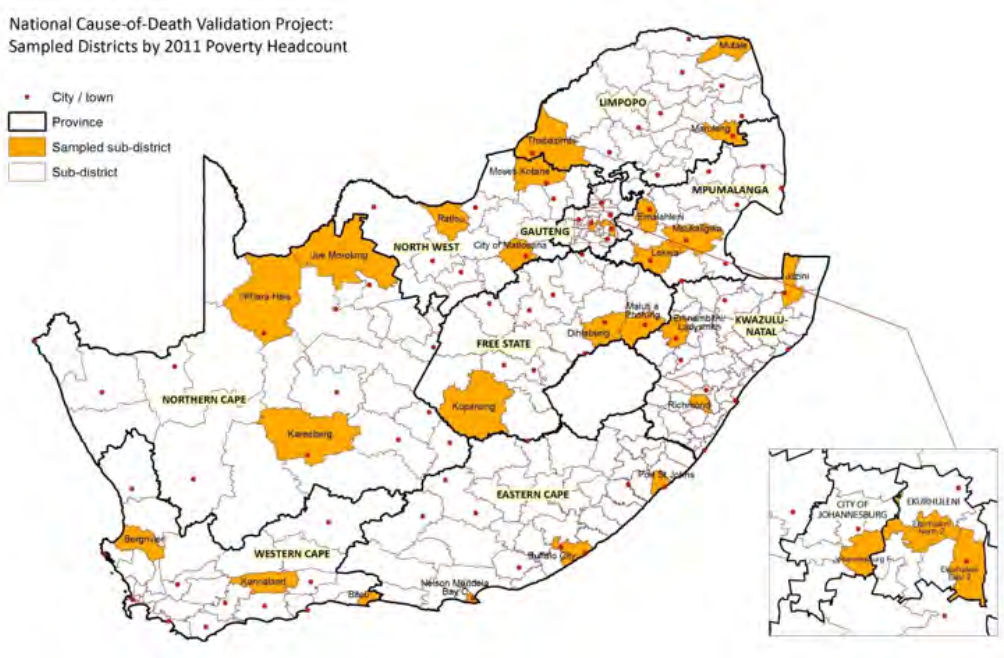


Figure 2. Map of selected health sub-districts and provincial boundaries, SA NCOD Validation Project 2017/18.

3.4 Sample size estimation

One objective of this study was the estimation of correction factors for cause-specific mortality fractions that can be applied to the CRVS cause-of-death data. This is different from the study by Rao *et al.*, (2007)[85] which aimed to directly estimate the cause of death profile i.e. the cause-specific mortality fractions from a sample. Thus, the sample size calculation for this project focused on generating adequate sample size for the correction factors for leading causes of death, namely, HIV, cerebrovascular disease, diabetes, and interpersonal violence which occur at different rates.

Sample size estimation needed to balance statistical considerations with logistical issues, and as such, was limited to a selection of leading causes of death from the South African National Burden of Disease Study 2012 (SA NBD 2012) cause list; the top 10 causes of death for 2012[86] are shown in Table 1. HIV was the leading cause of death and according to modeled estimates from the SA NBD 2012, CRVS data underestimated deaths due to HIV by approximately 80%. Thus, the correction factor for HIV was considered the main statistical measure to be estimated from the study and was anticipated to be of the order of 80%. However, to facilitate a standard calculation, it was assumed that the correction factor for any cause will be 50% as a worst-case scenario based on the variance of the binomial distribution.

Table 1: Top 10 causes of death in South Africa, SA NBD 2012.

	Cause of death	Number	%
1	**HIV/AIDS**	153,661	29.1
2	**Cerebrovascular disease**	39,830	7.3
3	**Lower respiratory infections**	25,977	6.4
4	**Ischemic heart disease**	24,969	5.6
5	**Tuberculosis**	23,817	5.4
6	**Diabetes mellitus**	18,894	5.1
7	**Hypertensive heart disease**	18,755	3.8
8	**Interpersonal violence**	18,741	3.1
9	**Road injuries**	17,597	2.9
10	**Diarrheal diseases**	16,349	2.8
Total		**528,947**	**100.0**

Source: Pillay-van Wyk et al. (2016).[86]

Table 2 shows the anticipated number of deaths for the selected causes (based on the 2012 cause-of-death profile). It then provides the adjusted number to create the impact of a design effect of 2 and assumes that data will only be collected from 85% of the deaths. Based on that number, Table 1 in Lwanga and Lemeshow (1991)[87] was used to determine the absolute error for an adjustment estimate of 50% (considered to be a worst-case scenario for estimation) based on the assumption of a binomial distribution. This was done for different combinations of N=10,000, 11,000 and 13,000 for the four selected causes of death to gauge the adequacy of the sample size and the value of increasing the sample size. Table 2 demonstrates that a sample size of 13,000 deaths would have produced 2-3% precision for HIV/AIDS; 4-5% for cerebrovascular disease; and 7% for diabetes mellitus and interpersonal violence, which were adequate for the objectives of the study. The anticipated breakdown of the total sample according to whether the death occurred in or out of hospital is shown in Figure 3, based upon current analyses of civil registration data.[55] It was expected that 1,040 of the cases would have been assessed in forensic pathology facilities.

Table 2: Sample size determination, SA NCOD Validation Project 2017/18.

Cause of death	Proportion of total deaths 2012	Impact of sample size											
		Number	Design effect*	Realization rate 85%	Precision (at 50%)	Number	Design effect*	Realization rate 85%	Precision (at 50%)	Number	Design effect*	Realization rate 85%	Precision (at 50%)
		10,000	5,000	4,250		11,000	5,500	4,675		13,000	6,500	5,525	
HIV/AIDS	29.1%	2,910	1,455	1,237	2-3%	3,201	1,601	1,360	2-3%	3,783	1,892	1,608	2-3%
Cerebro-vascular disease	7.5%	750	375	319	5-6%	825	413	351	5-6%	975	488	414	4-5%
Diabetes Mellitus	3.6%	360	180	153	8%	396	198	168	7-8%	468	234	199	7%
Interpersonal violence	3.5%	350	175	149	8%	385	193	164	7-8%	455	228	193	7%

* *Effective sample size if design effect = 2*

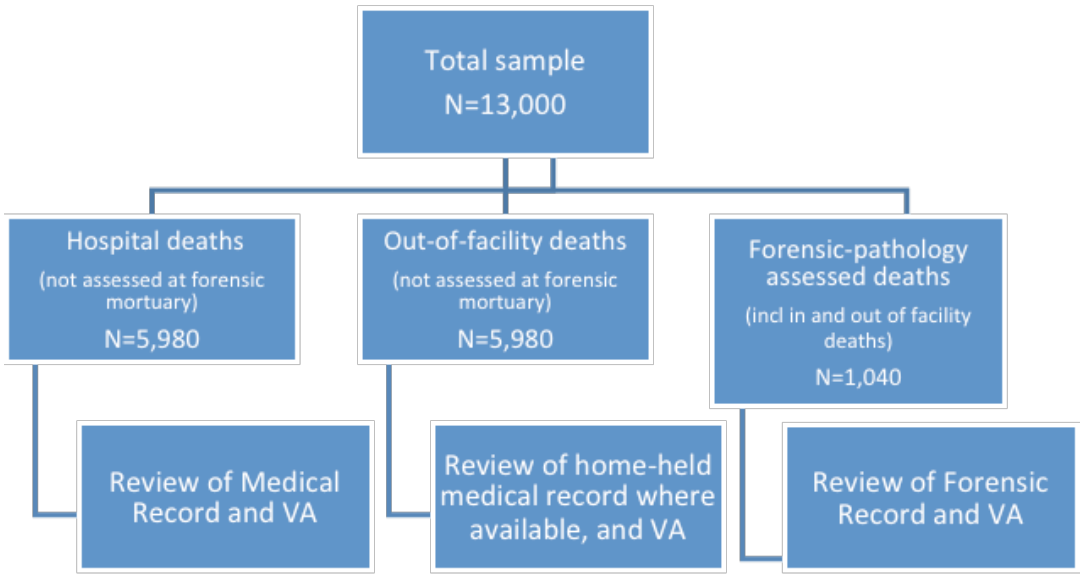


Figure 3: Graphical presentation of the sampling plan, SA NCOD Validation Project 2017/18.

3.5 Revised sample

By the end of the first 6 weeks of phase 1, it was recognised that recruitment of next of kin/carer participants for the study was extremely challenging. Different strategies were attempted to recruit more participants, but it became necessary to extend the study period beyond the original 3 months due to the small numbers recruited. At the end of the 8.5-month recruitment period, just over 6,000 next of kin/carers of decedents had agreed to be contacted to participate in the study through a verbal autopsy. This constituted less than half of the sample size of verbal autopsies determined to provide an adequate sampling error in the agreement and correction factor.

The sample plan as per protocol was to collect medical and forensic pathology records for the decedents for whom next of kin had consented. Given the challenges of the recruitment and the low recruitment, the protocol was amended to increase the sample size of the decedents who died in a health facility or were referred to forensic pathology services and ensure an adequate sample size for the respective sub-objectives of the validation (Annexure 8.1 Objective 1a and Objective 1c) as these data would provide the highest level of information for the validation and would reduce the sampling error of the sub-objective correction factor (Annexure 8.1 Objective 2a) and to some extent compensate for the loss of evidence from the verbal autopsies.

For the amended protocol it was anticipated that records for 16,730 deaths out of a total 33,313 deaths would be collected in the extended recruitment period by assuming that 45% of deaths would occur in health facilities and an 80% realisation would be achieved. Based on the sample size calculations shown in Table 2, it was concluded that this would be sufficient to meet the statistical requirement for the comparison of cause-of-death data from medical records with CRVS in the first sub-objective of the validation (Annexure 8.1 Objective 1a). Sampling strategies were considered to reduce the number of medical records to 13,000 but it was decided to keep the field team in each area slightly longer for the extra data collection and avoid a complicated sampling strategy to be implemented by the field team which would be difficult to monitor. Medical records were collected for all deaths that occurred in the facilities in the 27 sampled sub-districts.

In the amended protocol it was noted that although the study will provide invaluable information about the implementation of verbal autopsies, there is a possibility of bias in the data collected for the second validation sub-objective (Annexure 8.1 Objective 1b). It was proposed that, in the analysis of the linked data, it would be necessary to investigate the pattern of non-response and explore the possibility of doing a post-survey weighting, based on the basic characteristics of the registered deaths that occurred in the sampled areas when calculating the correction factors.

3.6 Data collection

Data collection was conducted in two phases. In the first phase, at the time of registering a death, next of kin were provided with information that the South African Medical Research Council (SAMRC) was undertaking a study about the causes of death and permission was sought for the SAMRC to contact them at a later stage to collect more information. In the second phase of the project, contact was made with the next of kin who had agreed to be contacted to arrange for a verbal autopsy interview. During this phase, medical and forensic pathology records were also collected for the deaths that had occurred during the study period. The overall data collection process is outlined in Figure 4.

3.6.1 Phase 1 of fieldwork – recruitment

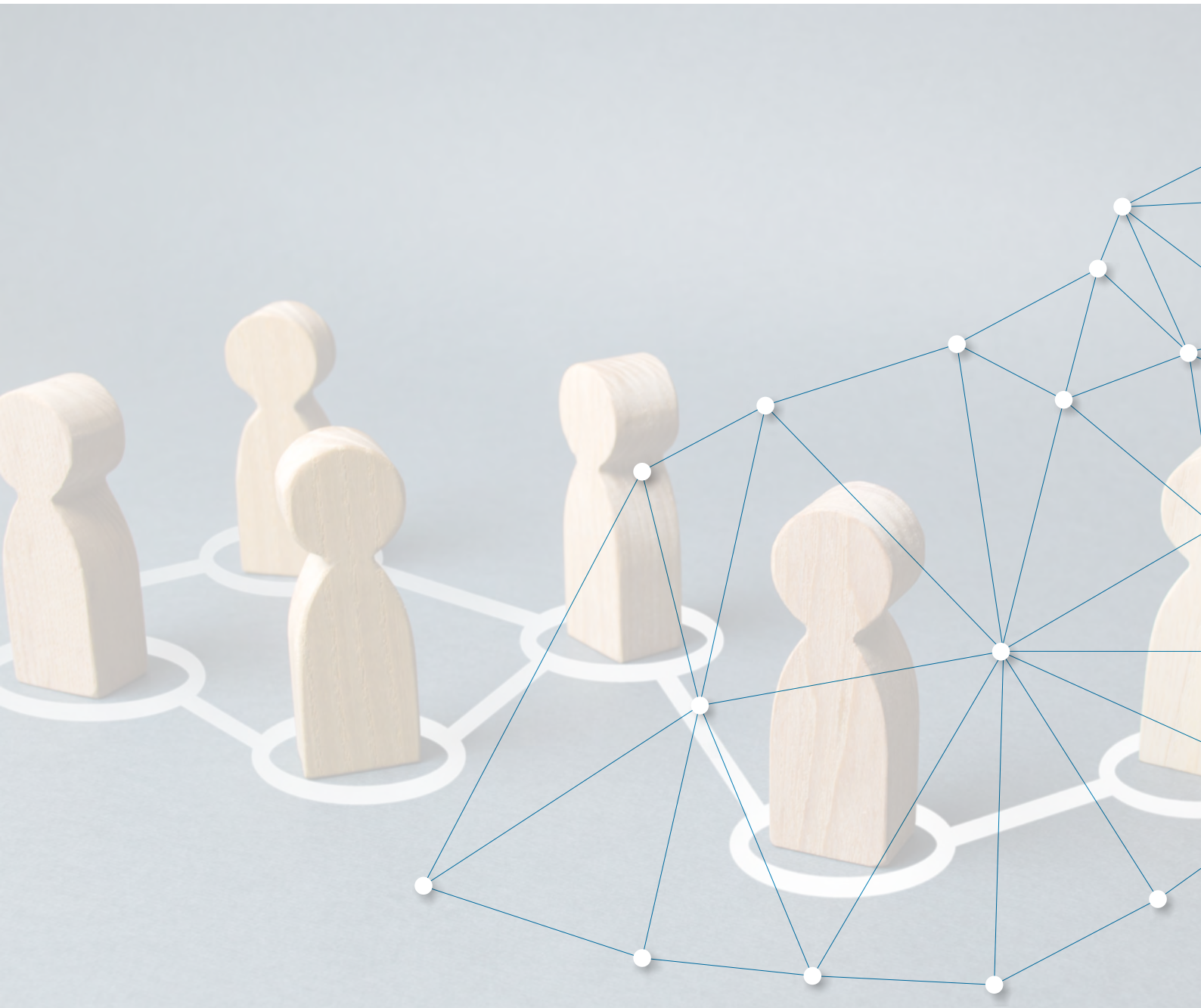
Next of kin of deceased were recruited through Funeral Practitioners (FPs), the usual frontline agent in the death registration process. As there is no complete list of funeral practitioners who operate in South Africa, it was necessary for the field workers to identify the funeral practitioners operating in the sampled areas as the first step and then approach them with an introduction to the project and a request for assistance. When funeral practitioners were recruited, they were provided with booklets to capture information about deceased persons and signed consents, as well as information sheets for the next of kin/informants and posters to alert the community to the project. Funeral practitioners were offered an amount of R35.00 (USDS2.5) for each recruitment as a token compensation for the time spent on this.

The method of identifying and recruiting FPs was pre-tested in the sub-district of Richmond where a favourable response was found. Field worker training was undertaken in August 2017 and recruitment of funeral practitioners started in September 2017. Prior to this, letters were sent to the associations of undertakers who were requested to alert their members in the sampled areas. Presentations about the project were made to some regional and one national meeting of undertakers.

During the pre-test, it was discovered that a number of funeral practitioners were not designated by Home Affairs. In such cases, the funeral practitioners are unable to register the death on behalf of the informant and the family needs to go to the Home Affairs office to register the death. The Department of Home Affairs indicated that they could not support the project to work with undesignated funeral practitioners, but they would allow the study to recruit "walk-ins" at the Home Affairs offices. After engaging with the Department of Home Affairs, it was agreed to present the project to the national

forum of branch managers in December 2016 and identify the method of collection of details of "walk-ins." Maps were provided to show the study area in terms of magisterial districts as well as municipal districts. In January 2017, the field team contacted provincial managers and office managers in the selected areas to set up recruitment in the offices. The fieldworkers made appointments to meet with officials in each office to explain the project and the recruitment process. Based on the number of "walk-ins" and the capacity of the official in each office, it was decided whether the project booklets could be left with the officials to record or whether an assistant from the field team needed to be assigned to the office to complete the paperwork. The Eastern Cape proved challenging as the branch manager had gone on leave without informing his staff of the arrangement. Through communication with the national office, it was possible launch collection in the Eastern Cape offices at a later date.

Funeral practitioners and Home Affairs officials were asked to identify all the deaths that occurred in the sample area during the study period and obtain the contact details for the next of kin who consented to be contacted by the research team. The study period was originally set from 1 September 2017 – November 2017 and the end date was later extended to 13 April 2018.



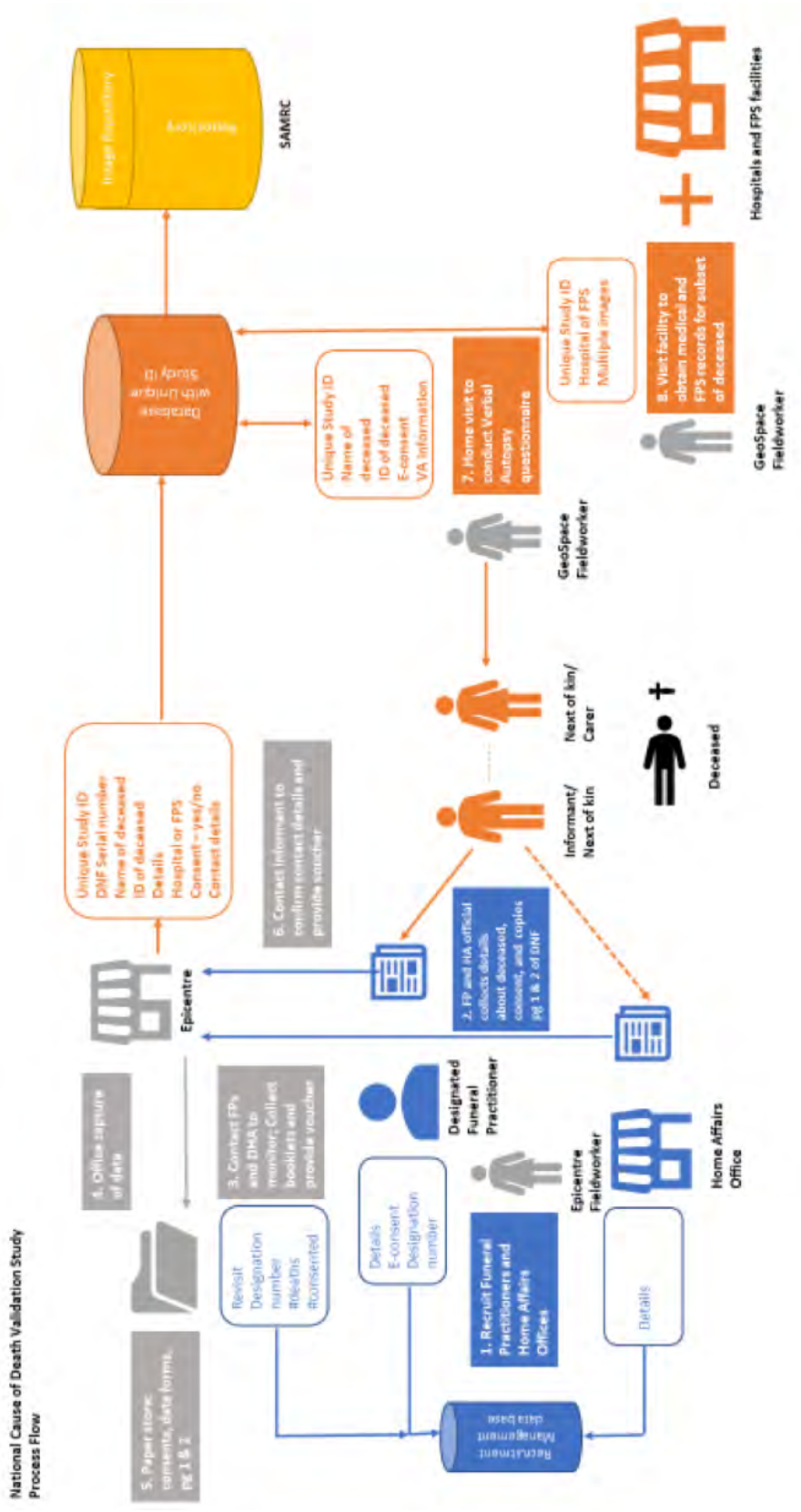


Figure 4: Schematic of data collection process, SA NCOD Validation Project 2017/18.

By the end of the recruitment phase spanning 1 September 2017 – 13 April 2018, a total of 353 funeral practitioners and 95 Home Affairs offices were engaged to recruit next of kin for the project (Table 3). A total of 9,730 next of kin were approached of which of 6,358 next of kin (65%) consented to be approached at a later stage by the project team. The total for each sub-district are shown in Table 3. Details of the deceased, including name and surname, identification number, date of birth and date of death together with the contact details and signed consent of the next of kin were hand-written in booklets of 20 cases. The booklets were returned to the field office where the data were captured into a spreadsheet to form the list of cases to be followed up in Phase 2.

Table 3: Total number of Funeral Parlors, Department of Home Affairs Offices and next of kin recruitment by health sub-district, SA NCOD Validation Project 2017/18.

Province	Health sub-district	Funeral Parlor Offices	Department of Home Affairs Offices	Next of kin		
				Approached	Recruited	Response rate
Eastern Cape	Buffalo City	40	4	497	285	57%
	Nelson Mandela Bay C	24	5	474	285	60%
	Port St Johns	6	2	137	125	91%
Free State	Dihlabeng	14	2	144	87	60%
	Kopanong	13	7	290	229	79%
	Maluti a Phofung	24	2	534	417	78%
Gauteng	Ekurhuleni east 2	14	1	546	329	60%
	Ekurhuleni north 2	17	3	259	89	34%
	Johannesburg F	14	9*	747	208	28%
KwaZulu-Natal	Emnambithi/Ladysmith	21	2	789	589	75%
	Jozini	3	6	656	340	52%
	Richmond	6	3	133	97	73%
Limpopo	Maruleng	1	4	192	154	80%
	Mutale	4	4	286	280	98%
	Thabazimbi	15	2	133	49	37%
Mpumalanga	Emalahleni	16	5	499	380	76%
	Lekwa	17	5	220	191	87%
	Msukaligwa	20	2	177	166	94%
Northern Cape	Joe Morolong	11	1	329	254	77%
	Kareeberg	3	2	33	23	70%
	Khara Hais	14	3	545	268	49%
North West	City of Matlosana	25	4	1,113	840	75%
	Moses Kotane	9	4	273	126	46%
	Ratlou	10	3	460	327	71%
Western Cape	Bergrivier	5	3	126	97	77%
	Bitou	4	3	55	44	80%
	Kannaland	3	4	83	79	95%
South Africa	**Total**	**353**	**95**	**9,730**	**6,358**	**65%**

* *1 in Johannesburg F; 8 in Soweto and other suburbs*

3.6.2 Phase 2 data collection

The WHO Verbal Autopsy 2016 questionnaires, the most up-to-date forms available, were selected for use in the study. Digital data collection tools were then developed using KoBoToolbox,[88] an open-source secure online/tablet platform set up by the Harvard Humanitarian Initiative for field-data collection in challenging environments. KoBoToolbox was selected because it is compatible with OpenDataKit (ODK)[89] allowing us to utilise the ODK versions of the three age-based VA questionnaires made available by the WHO. These questionnaires were translated by a registered company into the nine official languages considered most likely to be needed in the sampled areas. These included isiXhosa, isiZulu, isiNdebele, Afrikaans, SePedi, SeSotho, SeTswana, TshiVenda and XiTsonga (available on request from pamela.groenewald@mrc.ac.za). An independent review of each translation was undertaken by a second translator and any discrepant translations or concerns were discussed and resolved between the two translators. The translations were incorporated as digital questionnaire options to choose from on the KoBoToolbox platform. Answers were supplied in English on tablets by trained bi- and multi-lingual fieldworkers with a proven good command of English.

Several rounds of testing the questionnaires on the tool were undertaken to ensure the correct skip patterns were implemented as well as translation concerns. Following these evaluations, it was decided that the open-ended question to the next of kin about the circumstances of the death, which enables the fieldworker to compile a written narrative, should be asked at the beginning of the interview as it would facilitate the conversation during the interview. The narrative provided by each next of kin was hand-written on a note pad and later scanned with the tablet and uploaded into KoBoToolbox as an embedded document with the unique study identity number assigned to the associated decedent's VA questionnaire. Similarly, where a death certificate or Road to health card was available for the decedent, and the respondent consented for the interviewer to scan them, these documents were scanned using the tablet and uploaded to KoBoToolbox as an embedded document.

A medical record checklist[i] was set up using KoBoToolbox to capture identification details (name, surname, national identity number, date of birth, and date of death) against a unique study identity number for deceased hospital patients identified to be eligible for inclusion in the study. Inclusion criteria included a date of death between 1 September 2017 and 13 April 2018, and the hospital being in the selected health sub-districts, or the deceased patient being resident in the selected sub-district. Similarly, a forensic pathology checklist[ii] was set up in KoBoToolbox for decedents from forensic pathology facilities identified to be eligible for inclusion in the study. Any records for decedents whose next of kin had consented to a verbal autopsy were included. In addition, given the poor sample realization in phase 1, all medical and forensic records that met the above-mentioned inclusion criteria were included in the study.

Eighty-four fieldworkers were trained from 24 July 2018 – 7 August 2018 in Pretoria. Fieldworker applicants were scored based upon a matrix of education and experience. Graduates were preferred, but matriculants with adequate fieldwork experience were accepted. Experience in fieldwork with digital instruments was ranked as important as education qualifications. Good spoken and written English was a requirement as was multilingualism in any of the South African official languages. Team leaders required a driver's license and older persons with maturity were preferred for this role. A minimum of 50% females was also a requirement for selection. Each fieldworker was given a tablet that was set up with the data collection tools and fieldworker manuals.[iii] All fieldworkers were trained to conduct verbal autopsies with next of kin, capture identifiers from medical and forensic records, de identify medical and forensic records and scan all records from the last admission before death for medical records and all forensic records relating to the scene of the injury, postmortem results and any laboratory test results. The training included lectures[iv] to introduce the project, terminology and definitions, an overview

i Medical record checklist
https://www.samrc.ac.za/sites/default/files/attachments/2020-08-03/MedicalRecordChecklist.pdf

ii Forensic record checklist
https://www.samrc.ac.za/sites/default/files/attachments/2020-08-03/ForensicRecordChecklist.pdf

iii Verbal autopsy fieldworker manual available on request from pamela.groenewald@mrc.ac.za

of the content of the verbal autopsy questionnaires followed by an input on communication techniques to engage with households and how to deal with grief and anger from relatives of the deceased. Inputs on the content and managing the data capture tools were given with sessions for practice and roleplay. Continuous assessments were undertaken during the training to identify issues that needed re-training. There were also sessions on team work and logistics. Fieldworkers were given an orientation to the medical and forensic records that they would be required to scan at hospitals and forensic mortuaries. They were instructed to scan only the medical records from the last admission before death at hospitals, and all the documents in the forensic folders with the exception of the documents detailing the handover of the body which have no relevance to the cause of death. They were also trained in how to use Clearscanner in the classroom prior to the practice on real records later in the training and provided good quality records that could be examined in the office (see example copies of medical record and forensic pathology record).[iv]

Following the fieldworker training, hands-on field practice was arranged in a Community Oriented Primary Care project site where a ward-based outreach team could arrange for the fieldworker to conduct a verbal autopsy interview in a home where there had been a death in the past 12 months. In addition, fieldworkers were taken in groups to a forensic pathology mortuary and a hospital to practice reviewing medical and forensic record folders, and anonymising and scanning the relevant sections using the ClearScanner application.[90]

Debriefings were undertaken, and the captured data reviewed to provide feedback to the fieldworkers. Final corrections were made to the questionnaires and translations and field practice was continued until the teams felt confident to start collecting data. During the field practice, a total of 252 households were approached and 231 verbal autopsy interviews were conducted. Some amendments were made to the maternal questions of the verbal autopsy to make the questions clearer. These changes were incorporated into the translations. A total of 52 medical records and 38 forensic records were anonymized and scanned for practice purposes. These cases were not included in the final dataset. Since space and time was limited, only team leaders took part in the scanning exercise. During the main study fieldworkers captured study decedent identifiers from the medical and forensic records including name, surname, date of birth, date of death and SA identity number in the medical and forensic checklists previously described.

Fieldwork began on the 16th August 2018 in the Gauteng area so that the field team headquarters (Geospace International) could monitor and provide support. Teams were deployed to the various provinces at the end of August and a google sheet that could be accessed by the team leads was used for online field scheduling and monitoring progress. Generally, a team comprising four fieldworkers would conduct the verbal autopsies and scan the medical and forensic records for a specific sub-district before moving on to another sub-district. In some cases where permission to access hospitals or forensic mortuaries was delayed, a different team might have returned to do the record scanning. Fieldwork concluded in March 2019. Airtime and electricity token provision to respondents was managed and monitored online on a daily basis. Hospital and Forensic Pathology Mortuary communication was done on a daily basis to gain approval and access to the selected facilities to collect data. Quality assurance was set up at GeoSpace headquarters with daily review of the verbal autopsy questionnaires and the hospital and forensic pathology records. The project team reviewed the data collected on KoBoToolbox on a weekly basis and any issues were discussed with the field team manager.

iv Available on request from pamela.groenewald@mrc.ac.za

3.7 Data Processing

3.7.1 Doctor review of verbal autopsies

We recruited medical doctors to participate in the study through an advertisement posted on the SAMRC website and shared with colleagues. The doctors were required to attend face-to-face training for 1 day and to successfully complete 3 home assignments and pass a competency test before they were offered a contract. The main aim of the training was to ensure that the doctors were competent in certifying deaths according to ICD-10 guidelines and were able to use the data capture tool, understand the WHO 2016 Verbal Autopsy tool, and interpret the verbal autopsy narrative and interview. A training manual,[v] a series of PowerPoint presentations[vi,vii] and class assignments to certify medical cause of death[viii] and review verbal autopsy examples[ix] were used during face to face training. Participants were required to complete a home assignment on medical certification of cause of death[x] as well as a review of verbal autopsy interviews[xi,xii] and a competency test.[xiii] Standard operating procedures[xiv] were developed and shared with the reviewers via the Microsoft Teams application.[91] This included technical SOPs for using KoBotools and KoBoCollect.[xiv]

There were 105 doctors who attended the training of whom 75 successfully completed the assignments and were appointed on services rendered contracts. Seventeen doctors resigned between March 2019 and November 2019 leaving 58 still conducting reviews at end November 2019. The majority of the doctors were doing the reviews after routine work hours.

Each verbal autopsy interview was allocated a unique study identity number (USID). The first digit represented the province, the second and third represented the subdistrict and the last four digits were a sequential numbering generated within each subdistrict for the consented next of kin. The verbal autopsy interview answers submitted by field workers were summarized in an Excel worksheet and the narratives for each interview were saved as pdf files, named by the USID, and grouped into batches of 40 verbal autopsies using Microsoft Teams. Each batch of verbal autopsy records were independently reviewed by two doctors who accessed the data in Teams and captured their record reviews in KoboCollect on an android tablet or KoBoToolbox using a personal computer, using the customized verbal autopsy record review form.[xv] Each doctor's verbal autopsy record review form captured a short summary of the case, information on HIV and TB, the manner of death and the sequence of medical conditions leading to the death as would be reflected on a certificate of cause of death according to ICD-10 guidelines. Reviewers also had space to provide feedback about each case, if desired.

v Training manual for medical doctors for reviewing verbal autopsy, medical and forensic records
https://www.samrc.ac.za/sites/default/files/attachments/2020-08-04/TrainingManualMedicalDoctors.pdf

vi Medical certification of cause of death training
https://www.samrc.ac.za/sites/default/files/attachments/2020-08-04/MedicalCertification.pdf

vii Verbal autopsy physician assessment training
https://www.samrc.ac.za/sites/default/files/attachments/2020-08-04/VerbalAutopsyPhysicianTraining.pdf

viii Medical cause of death certification classroom assignment
https://www.samrc.ac.za/sites/default/files/attachments/2020-08-04/MedicalCertificationCause%20DeathClassAssignment.pdf

ix Verbal autopsy review classroom assignment
https://www.samrc.ac.za/sites/default/files/attachments/2020-03-03/VAreviewClassAssignment.pdf

x Medical cause of death certification home assignment
https://www.samrc.ac.za/sites/default/files/attachments/2020-08-05/MedicalCertificationCauseDeathHomeAssignment.pdf

xi Verbal autopsy review narratives home assignment
https://www.samrc.ac.za/sites/default/files/attachments/2020-08-05/VerbalAutopsyReviewHomeAssigment.pdf

xii Verbal autopsy review anon data home assignment
http://www.samrc.ac.za/sites/default/files/attachments/2020-02-17/VAdata_home%20assignment.xlsx

xiii Clinician reviewer competency test
https://www.samrc.ac.za/sites/default/files/attachments/2020-08-05/ClinicalReviewerCompetencyTest.pdf

xiv Consolidated doctor reviewer technical support
https://www.samrc.ac.za/sites/default/files/attachments/2020-08-04/DoctorReviewerTechnicalSupport.pdf

xv Verbal autopsy review data capture form
https://www.samrc.ac.za/sites/default/files/attachments/2019-12-12/Verbal%20Autopsy%20review%20data%20capture%20form.pdf

After the training by one of the co-principal investigators, a team of five medical doctors who demonstrated excellent competency in medical certification were recruited to provide quality assurance of the reviews. The quality assurance reviewers were allocated batches of verbal autopsy reviews to assess. Each review was assessed and where the two independent reviews had different underlying cause of death the reviewers were informed and asked to come to consensus on the casual sequence in Part 1 of the medical certificate of cause of death. Where they could not come to consensus, they were asked to inform the quality assurance reviewer who either sided with one of the reviewers or brought it to a panel review of the whole Quality Assurance (QA) panel. Once the review of the batch was completed it was signed off by the quality assurance reviewer.

A neonatal expert panel was set up to clarify the ICD-guidelines on certification of the causes of perinatal deaths and to review cases where consensus could not be reached by independent reviewers. A specialist obstetrician reviewed all maternal deaths independently for an additional sensitivity analysis.

During the training of doctors in medical certification of cause of death it became apparent that additional training in medical certification of cause of death for medical doctors in South Africa would be useful beyond this study. In order to reach as many medical doctors as possible it was decided to use the project training materials to develop an online course on medical certification of cause of death (www.deathcertification.org).

3.7.2 Doctor review of medical records

The majority of the medical reviewers recruited for the verbal autopsy reviews also conducted the medical record reviews (30/49). Additional training materials[xvi] and five test medical records[xvii] were provided to orientate the existing reviewers to the medical record reviews and the medical record review data capture form.[xviii] Feedback was provided on the reviews for the test medical records. Only reviewers who provided reviews of acceptable standard were asked to continue with medical record reviews.

Additional recruitment was undertaken to assist with the medical record reviews. Face-to-face training on ICD-10 guidelines on medical certification of cause of death[v] and the KoBoToolbox medical record review data capture form[xviii] was provided to new reviewers. The new reviewers were required to successfully complete the medical certification home assignment[x] and competency test[xiii] as well as five test medical records[xvii] before they were offered a contract. An additional 16 reviewers were recruited to do medical record reviews.

Medical records consisted of pdf files of images of the decedent's medical records from their last admission to hospital before death. Batches of 40 medical records were prepared and allocated to reviewers using Microsoft Teams. Reviewers viewed the records on their laptops and they captured the information extracted using a medical record review form in KoboCollect on an android tablet. Only one review was conducted per medical record.

The medical record review form filled out by reviewers captured a short summary of the decedent's medical history, information on TB and HIV status, manner of death and the sequence of conditions leading to death according to the format of the certificate of cause of death along with an indication of the level of certainty for the causes listed in Part 1.

xvi Guideline for reviewers: Medical certification of death
https://www.samrc.ac.za/sites/default/files/attachments/2020-08-04/GuidelinesMedicalCertification.pdf

xvii Available upon request from pamela.groenewald@mrc.ac.za

xviii Medical record review data capture form
https://www.samrc.ac.za/sites/default/files/attachments/2019-12-12/Medical%20record%20review%20data%20capture%20form.pdf

The same team of five quality assurance reviewers plus an additional recruit who excelled in certification of cause of death conducted the quality assurance. All cases with an unknown underlying cause of death were reviewed in detail against the medical record to ensure that no information had been missed. Where additional information was found by the quality assurance reviewer, the final underlying cause was decided by consensus in among the panel. In addition, 4 records from each batch were randomly sampled for quality assurance to check whether the quality assurance reviewer agreed with the underlying cause selected by the medical reviewer. If the quality assurance reviewer's opinion on cause of death differed with the medical reviewer for two or more records, the whole batch was assessed, and feedback was given to the medical reviewer. Differences of opinion or in cases where it was difficult to identify a cause of death, the case was discussed in a weekly panel review consisting of the quality assurance reviewers and the co-principal investigator.

3.7.3 Forensic pathologist review of FPS records

A total of 11 doctors were trained to perform forensic record reviews, including three doctors who participated in the verbal autopsy reviews and an additional eight doctors who were recruited solely for the forensic record review. Once the three doctors who participated in the verbal autopsy reviews completed their verbal autopsy reviews, they were oriented to the Forensic record review and conducted 2-5 forensic record reviews prior to being allocated batches of 40 forensic pathology records for review in a similar manner to the Medical records.

The additional 7 reviewers who were recruited for forensic record reviews received face-to-face training on ICD-10 guidelines on medical certification of cause of death and the KoBoToolbox forensic record review data capture form. They were also required to conduct 2 – 5 forensic record reviews prior to being allocated batches to review. Whilst all those trained were eligible to review records, only 4 went on to perform reviews, mainly due to work commitments.

The forensic record review form[xix] in KoBoToolbox completed by the reviewer captured a short summary of the decedent's case history, information on HIV and TB status, manner of death and the sequence of conditions leading to death according to the format of the certificate of cause of death along with an indication of the quality of the forensic records and level of certainty for the underlying cause of death.

A small team of quality assurance reviewers reviewed all the forensic records to ensure that the certification of cause of death included the circumstances of the death as well as the manner of death. All records with unknown underlying cause of death were reviewed against the forensic records to ensure that no information had been missed. Where necessary, these cases were discussed with the reviewer to reach consensus on manner and circumstances of the death. Where consensus could not be reached between the original reviewer and the quality assurance reviewer, the case was referred to the panel of quality assurance reviewers for discussion and a decision on the underlying cause.

3.7.4 Coding cause of death

All cause of death coding has been performed by the research team after field work had been completed and the data sets cleaned. The InterVA-5 tool has been used for automated selection of the most probable underlying cause of death, based on the responses to the verbal autopsy questions.[92] This is based on the responses to the Verbal Autopsy interview and does not take the narrative into account. The list of causes is restricted to the 64 causes listed in the WHO 2016 VA cause of death list from the WHO 2016 VA instrument. This tool has included an innovative categorisation based on selected questions to provide information about the social and health circumstances of death. The COMCAT[93] categories will assist in contextualizing the determinants of the death in addition to the medical conditions by identifying whether health systems issues, care seeking behavior, resources etc. contributed to the death. The verbal autopsy survey data was downloaded from Kobotools by the SAMRC research team and formatted for input into InterVA-5 before processing by InterVA-5.

xix Forensic record review data capture form
https://www.samrc.ac.za/sites/default/files/attachments/2019-12-12/Forensic%20record%20review%20data%20capture%20form.pdf

ICD-10 coding of the multiple causes of death and the underlying causes of death from the medical certificates of cause of death produced by the doctor reviews of verbal autopsy interviews and narratives, was done using Iris automated software.[94] The verbal autopsy doctor review data was downloaded from Kobotools by the SAMRC research team. From this data, the Access data files required for Iris were prepared for batch processing. The dictionary for medical terms (text) to ICD-10 codes developed for the Western Cape local mortality surveillance system was updated and used. Rejects were manually coded by two researchers and a Co-principal investigator who had training in ICD-10 coding.

3.8 Data management, cleaning and analysis

3.8.1 Data management

Phase 1 data were collected using the Mobenzi platform.[95] In Phase 2, data were collected using KoBoToolbox using access controlled android tablets. Both the Mobenzi and KoBoToolbox servers provided excellent security and reliability including physical access control, and an online protection through a firewall to protect against hacking and viruses.

At the sampled hospitals and forensic pathology mortuaries, fieldworkers captured personal identifiers from relevant medical and forensic records using KoBoToolbox data collection forms for a medical record checklist and a forensic record checklist and issued a unique study identifier if the decedent did not already have one. To ensure confidentiality. Pages from medical and forensic records were anonymized by covering any patient identifiers with sticky notes and labelled with the assigned unique study identification number. Imaging was done by fieldworkers using a high-definition camera software application, ClearScanner,[90] using the android tablets. The collected images were stored on the access-controlled device and uploaded daily to the secure access-controlled Dropbox for Business[96] folder.

In compliance with SAMRC Information Technology policy, images of anonymised verbal autopsy narratives and data medical and forensic records were stored on Microsoft Teams for access by the medical reviewers. This allowed for restricted access and provided a secure platform for data storage. Medical reviewers accessed relevant records on Microsoft Teams on their laptops and captured record review data in KoBoToolbox data collection forms that had been installed on their password protected android tablets. KoboCollect software installed on the tablets collected all data submitted into KoBoToolbox forms and automatically uploaded data to a secure server based at Harvard University from which the data could be downloaded by the research team at SAMRC. The anonymized review data are stored on a secure share drive on the SAMRC server. The verbal autopsy review data were coded using Iris automated software as described in the previous section on data processing.

Data access was restricted to authorized users only, with a full audit trail maintained to guarantee data integrity. User access was limited to the information pertinent to that user. United States Centers for Disease Control and Prevention (CDC) staff were not involved in data collection and did not have access to participants' identifying information. Once the study was completed, a backup of the VA questionnaire data and the patient records data, excluding the identifying information, was archived, and the identifying information deleted from the server of the service provider. Funeral practitioner booklets will be stored in a locked facility at the SAMRC for five years thereafter they will be destroyed. Electronic records will be retained for five years on the SAMRC secure server.

3.8.2 Data cleaning

The identifiers (including names, South African identity number, date of birth, date of death and gender) from the three datasets (verbal autopsies, the medical record checklist, and the forensic pathology checklist) were merged on the unique study identifier to create a consolidated Master List of the decedents in the study. We checked that the South African national identity numbers (SA ID) numbers were valid. Invalid SA ID numbers were identified through an algorithm and the last digit (13th), corrected according to the sequence of the first 12 digits.[97] In the cases where the first 6 digits of the invalid SA ID numbers did not reflect the date of birth, these were corrected accordingly and again verified using the algorithm. The corrected SAID numbers were then linked to the Rapid Mortality Surveillance database[98] to verify that the death had been registered. Linkage to the Rapid Mortality Surveillance database was done on date of birth, date of death, sex and province for records that did not have ID numbers. When a definite match was found, the SA ID number was included in the consolidated Master List.

The identification of duplicate records of the same decedent was conducted on SA ID number as well as on the combination of date of birth and date of death. In cases where duplicates were identified across any of the 3 data sources, exact cases were identified and dropped from the Master List and cases with the same USID number (but that were not the same decedent) which arose from the algorithm that we applied during data collection to cater for simultaneous data capture from multiple facilities, were identified and a new unique USID number was allocated.

The ICD-10 codes for the underlying cause of death from doctor-certified and Iris coded VAs were run through the ANACONDA tool,[99] to ensure that no biologically implausible causes had been assigned. Six cases were identified as having biologically implausible causes, based on sex or age of the decedent. The underlying cause was modified for three of the cases to ensure plausibility. However, the panel considered that cerebral palsy was biologically plausible for the neonatal period and no change was made for the three such cases.

3.8.3 Data analysis

The workflow of the verbal autopsy interview data is shown in Figure 6. After the data collection was completed, the data was cleaned and the responses to the verbal autopsy interviews recoded to suit the InterVA-5 package. This was used to generate up to five probable causes of death classified according to the WHO 2016 verbal autopsy cause of death list with an associated probability. The cause with the highest probability has been used in this analysis and reported according to the WHO 2016 cause of death list for verbal autopsy (64 causes), and the burden of disease 3 broad cause groups with an additional category for HIV/AIDS and TB.

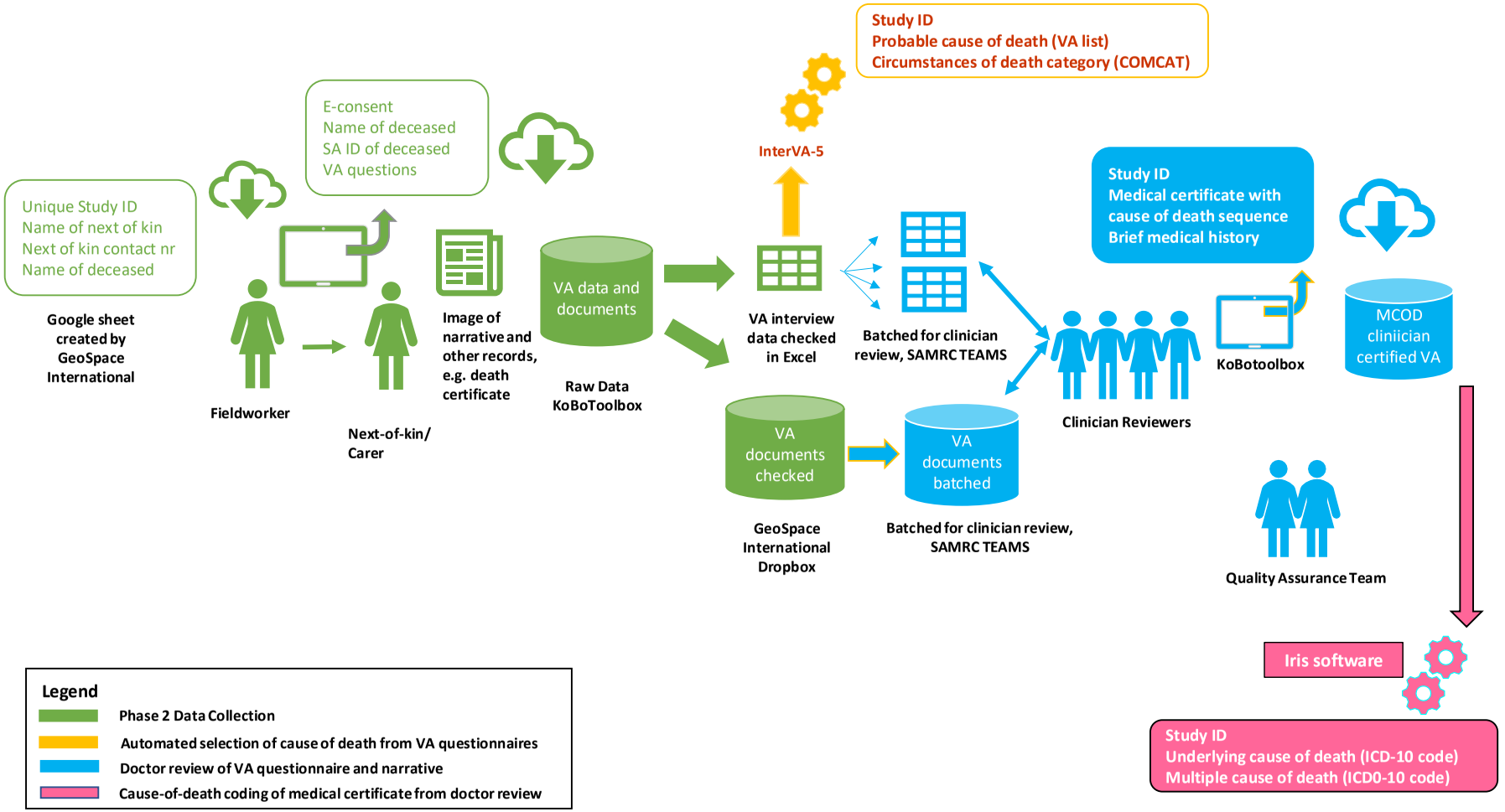


Figure 5: Verbal autopsy interview data workflow, SA NCOD Validation Project 2017/18.

The data and records from the verbal autopsy interviews were batched and distributed to doctors for reviews. The medical certificate data provided by the doctors were quality assured ensuring a consistent set of information for each death. The physician review data was coded using Iris which provided underlying cause of death by 4-digit ICD-10 codes. The causes were aggregated to the following groups: ICD-10 Chapters; the WHO 2016 cause of death list for verbal autopsy (64 causes), and the burden of disease 3 broad cause groups with an additional category for HIV/AIDS and TB. The underlying causes of death were assessed using the quality criteria developed by an expert group convened by the Bloomberg Philanthropies Data for Health Initiative and the Civil Registration and Vital Statistics Improvement project of the University of Melbourne in 2017.[100] Five categories of "unusable" codes were identified including immediate causes of death (e.g. Disseminated intravascular coagulation [defibrination syndrome]), Impossible as underlying causes of death (e.g. Viral warts), Insufficiently specified causes within ICD chapter (e.g. Cancer with unknown primary site), Intermediate causes of death (e.g. Other cardiac arrhythmias) and Symptoms, signs and ill-defined conditions (e.g. Headache, Other abnormal findings of blood chemistry). For the deaths due to HIV and/or TB, the short medical history provided by the doctor (based on the verbal autopsy interview and the narrative) was examined for any mention of defaulting treatment.

Descriptive statistics of the basic characteristics of the deaths with verbal autopsies including median age, and proportions, and 95% confidence intervals were calculated for sex, province, and population group using the svyset command in Stata IC/14.2 (StataCorp College Station, TX, USA) to allow for the stratification of the sample and Excel for Microsoft Office 365 ProPlus Version 1902 (Build 11328.20480 Click-to-run). The population group was based on the classification used by Statistics South Africa (black Africans, whites, Asians, and coloureds) since inequalities in health persist.

In the final section of the results, the Statistics South Africa 2016 data as well as the VA doctor review data and InterVA data were aggregated into the South African 4 broad cause groups and compared.

The anonymised data set will be made available coincident with the publication of papers reporting the findings of this study. The final anonymized dataset will be archived and stored with metadata for 20 years in a data repository at the SAMRC.

3.9 Ethical consideration and permissions

3.9.1 Information from informants and patient records

The primary subjects whose information was used to achieve the objectives of the project are deceased individuals. To gain access to relevant information of the decedents, the project required interaction with the next of kin or carer of the decedent through individual face-to-face verbal autopsy interviews. Since the Protection of Personal Information Act (POPI), 2013[101] precludes the Department of Home Affairs or the Department of Health from sharing the personal information of the next of kin of decedents without their consent, we had to work through an intermediary who could request consent from the next of kin for their contact details to be shared with the project. Funeral practitioners and Home Affairs officials were identified as intermediaries for the project. In addition, data were collected through the review of patient records held in the health information systems of public hospitals and forensic pathology mortuaries.

To ensure the highest level of respect for individual decedents and the informants who answered the verbal autopsy interviews, ethical clearance for research involving human participants obtained from the SAMRC Ethics Committee, and Health Research Ethics Committees required at provincial and health facilities as required. The protocol was also reviewed in accordance with the Centers for Disease Control and Prevention (CDC) human research protection procedures and was determined to be research, but CDC investigators did not interact with human subjects or have access to identifiable data or specimens for research purposes.

Another major ethical consideration in the project referred to the confidentiality of information from medical and forensic records and informants. As far as possible, anonymized decedent data were used as input to the project. Strict confidentiality measures were adhered to with regard to the protection of information obtained from medical and forensic records.

3.9.2 Informed consent

Informed consent was obtained from participating funeral practitioners and Home Affairs officials when they recruited informants as outlined in the methods. An information booklet was provided to funeral practitioners and Home Affairs officials, and their electronic signature was obtained along with their details (name and designation number and contact details of funeral company) the time of their training. Funeral undertakers and Home Affairs official obtained permission from the next of kin for the project to contact them to arrange an interview from the person/s who reported the death to the funeral parlor. The next of kin were informed about the aim of the project by the undertaker, using an information sheet which explained the aim, methods, and envisaged outcomes of the project. The undertakers and Home Affairs officials requested signed consent on the information sheet, and contact details for the next of kin, and recorded these on the Funeral Practitioner Death Register Form. Non-literate informants were asked to identify a person that they would be comfortable with to serve as an impartial witness to support them through the consent process, after which the volunteer provided a cross (X) to indicate consent. These forms were signed by the witness.

In the second phase of fieldwork, the fieldworker explained the aim, methods, and envisaged outcomes of the project to the family again to ensure that the participant was fully informed before signing consent for the verbal autopsy interview on the tablet. Provision was made for illiterate respondents to identify an impartial witness to support them through the consent process, after which the volunteer drew a cross with his/her finger on the tablet, and a witness signed a document to declare that the cross belonged to the particular volunteer.

3.9.3 Permission

Permission to access information of decedents from medical and forensic records at public hospitals and forensic autopsy facilities was obtained from the national, provincial and district health departments as well as individual facilities. Permission to access forensic pathology records in KwaZulu Natal could not be secured. The protocol was presented to the National Forensic Pathology Services Committee to obtain their support.

3.9.4 Confidentiality

The importance of confidentiality was explained to all fieldworkers during training and all other project staff including field supervisors, researchers, quality assurance staff, data managers, and research/administrative staff, information technology support staff and the doctors undertaking the reviews. All project staff were required to sign a confidentiality agreement to handle all project data ethically and confidentially.

Researchers and field workers had access to individual patient records in multiple formats, including individual paper-based or electronic in-patient records, and paper-based or electronic registers which include entries for individual patients and verbal autopsy interviews. Individual decedent data were de-identified, as described in the Data Collection section once a unique study ID was allocated. Data provided to the doctors to review were anonymised and identified through a unique study ID number. A master index was created with names and other identifiers and the study ID number which will be used for the data linkage in the final stage of analysis. The master index file, with restricted access, comprises the identifiers of all the deceased but has no data concerning the cause of death.

Results produced from the project will be in aggregate form and will not be able to be traced back to individual decedents.

3.9.5 Reimbursement

The funeral parlor directors were reimbursed for their time with a voucher of R35.00 (USD2.5) per next of kin recruited. Record clerks in hospitals and forensic mortuaries were compensated for their support with a small edible gift of biscuits, valued at R35.00 (USD2.5) per institution. A token of airtime or electricity worth R35.00 (USD2.5) was given to the interviewee at the time of the interview.

3.9.6 Potential risks and benefits

This study was reviewed and determined to involve minimal risk to participants. However, since participants were asked questions about the circumstances around the death of a family member which may have caused emotional distress, verbal autopsy interviewers were provided with skills to understand grieving, not to over-identify with interviewees, and to show sensitivity in questioning and probing. They were given skills to address any potential stress or discomfort that may result from study participation, and to help make participants manage their grief.

In cases where potentially unlawful acts were disclosed during the interviews, or instances of suspected child abuse, project staff were required to report this to the relevant authorities for further investigation, under section 110 of the Children's Amendment Act, No. 41 of 2007. However, where the interviewee expressed the opinion that the death was intentional, either due to self-harm or homicide, the interviewer informed them that, if they suspected that the death was caused by unnatural causes, they are legally required to report it to the police, under the Inquests Act, No. 58 of 1959. Since the interviewer would only have hearsay evidence, they were not required to report this themselves.

As part of the informed consent procedure, all potential participants were informed that they that they could withdraw from the study at any time with no consequences. Field staff assured interviewees that all responses and information would remain confidential.

Field staff were trained to refer participants for counselling support, if necessary. Project staff identified referral networks and social workers linked to clinics in the study areas prior to commencing fieldwork. All participants, irrespective of demonstration of pain, stress or trauma, were informed of available support services in their communities, should they have felt the need for further support.

Benefits include improved quality of cause of death data for health policy makers, as well the strengthening of research and analytic capacity through the methods and staff development for the project, but also via consultation with and technical inputs by expert co-investigators and technical advisors working with the research team.

All adverse events were reported to the Principal Investigator and an action plan implemented and reported to the SAMRC Ethics Committee. Five cases were referred to the principal investigator, of which three were confirming that the study was genuine, one wanted to report that her child's murderer had not yet been charged, and one had gone to the hospital where her child had died to request her medical records. A staff member from the hospital had called to investigate this. The interviewee had misunderstood the verbal autopsy question about whether she had any medical records for the decedent. The interviewer reported that the interviewee had been emotional and a little confused during the interview – she called her back to explain and resolve the situation. None of these cases required formal intervention or retraining of study staff.

3.9.7 Data linkage with national CRVS mortality dataset

The national civil registration and vital statistics dataset to be used for data-linkage in Phase 4 was processed, coded, and analysed at Stats SA, the official national statistical agency. The notification and registration of deaths in South Africa are mandated by the Births and Deaths Registration Act of 1992,[34] and the capturing, processing and dissemination of cause-of-death data from death notification forms are governed by the Statistics Act of 1999. Only 2017 civil registration and vital statistics data will be available in 2020 so this linkage exercise will be restricted to the 2017 study data.

The linkage exercise will include electronic handling of electronic records with identifiable information of the decedents. Physical and electronic safety measures will be taken to ensure confidentiality of the identity, and protection of private information of all decedents. The matching will be done in compliance with the legal requirements of the South African Statistics Act,[102] and according to standard in-house procedures of ensuring confidentiality and protection of the data.

The Act requires that the confidentiality of the identity of the deceased, and the information provided by informants reporting the death, must be protected, and that any results of the analysis of statistical information may not be disseminated in a way which is likely to enable the identification of an individual.[102] The linkage will be conducted within the Head Office premises of Stats SA in Pretoria under supervision of the Director: Births and Deaths. The data, with ID numbers, are stored on site in the Stats SA SAS library for mortality and causes of death. Access to this library is restricted to authorized staff members only, including the cause-of-death processing team, and the Births and Deaths Directorate. Access is only granted on site at the Head Office, and access rights for an experienced cause-of-death analyst and intern will be requested from the Director: Births and Deaths.

The final analysis dataset will be de-identified in a manner that no re-identification of any decedent will be possible.

4. RESULTS

4.1 Sample realisation

Data obtained from the Department of Home Affairs based on place of death within the sampled sub-districts indicated that there were 36,976 deaths registered in the study period 1 September 2017 – 13 April 2018. The provincial breakdown is shown in Table 4 together with the realisation of the sample. Across the sampled sub-districts, the number of verbal autopsies conducted accounted for 15% of the number of deaths registered during the study period. A total of 18,630 medical records and 5,915 forensic pathology records were collected. Due to file syncing issues with the uploading of scanned documents from the field, 1,005 medical records and 173 forensic pathology records were lost. Although the fieldworkers had scanned them, they did not get loaded into the data repository. These accounted for 5.4% and 2.1% of the respective samples, leaving 17,725 medical records and 5,742 forensic pathology records for review. The final numbers of records available for analysis are shown in Table 4. Overall, data were obtained from at least one source (i.e. verbal autopsy, medical record or forensic pathology) for 72% of the deaths registered during the study period, which ranged from 46% in Gauteng to over-representation of 304% in Limpopo and 434% in the Western Cape. In these provinces, it is clear that it is clear that deaths from neighboring areas were included in the sample.

The number of decedents in the realised sample according to the source of data is shown in Table 5. The majority of the deaths had data from a medical record only (59.4%) and a further 20.4% had data from a forensic pathology record only.

There were 3,291 (12.4%) decedents for which a verbal autopsy interview alone was conducted.

Table 4: Numbers of registered deaths, numbers of deaths for which data were collected during study period, and proportion of deaths with data by province, SA NCOD Validation Project 2017/18.

Province	Registered deaths	Decedents for whom verbal autopsy was completed	Ratio number of deaths with verbal autopsy relative to deaths registered (%)	Deaths in which medical records were collected	Deaths in which forensic pathology reports were collected	Deaths with at least one data collection source	Ratio of number of deaths with at least one data collection source relative to deaths registered (%)
Eastern Cape	12,960	575	4.4	4,550	1,318	6,147	47.4
Free State	3,442	656	19.1	1,504	391	2,313	67.2
Gauteng	9,371	463	4.9	3,139	809	4,262	45.5
KwaZulu Natal	2,240	890	39.7	1,674	0	2,214	98.8
Limpopo	635	377	59.4	1,090	606	1,918	302.0
Mpumalanga	3,114	610	19.6	1,282	721	2,336	75.0
Northern Cape	1,131	506	44.7	849	570	1,667	147.4
North West	3,570	1,108	31.0	1,870	853	3,449	96.6
Western Cape	513	202	39.4	1,667	474	2,208	430.4
Total	**36,976**	**5,387**	**14.6**	**17,625**	**5,742**	**26,514**	**71.7**

Table 5: Number and proportion of decedents in sample according to source of data, SA NCOD Validation Project, 2017/18.

Data collected	Number	%
Verbal autopsy interview, hospital and forensic pathology record	18	0.1
Verbal autopsy interview and hospital record	1,809	6.8
Verbal autopsy interview and forensic pathology record	298	1.1
Verbal autopsy interview only	3,262	12.3
Hospital and forensic pathology record	97	0.4
Hospital record only	15,701	59.2
Forensic pathology record only	5,329	20.1
Total	**26,514**	**100.0**

While the apparent overall response rate of the sample was high (72%), some of the deaths for which verbal autopsy data were collected were not actually part of the target study population. The geographic location of the address where the verbal autopsy interviews were conducted is known and was often outside the sampled sub-districts (Figure 6). In some instances, the next of kin/carer was resident outside of the sampled area by the time of interview. These were included in the sample. However, it was found that 58.9% of the 4,928 deaths with verbal autopsy and known SA ID number merged with the 36,976 registered deaths from the Department of Home Affairs, indicating that they did not occur in the sampled areas. This ranged by sub-district from 4% in the Western Cape sub-district of Kannaland and the Limpopo sub-district of Mutale to 94% in the Free State sub-district of Maluti a Phofung. Some areas were particularly difficult to delineates as the Department of Health operates within the municipal boundaries while the Department of Home Affairs operates within the magisterial district boundaries. Furthermore, the municipality of Mutale is currently undergoing extensive changes to its boundaries. The place of death for the cases with a medical record or forensic pathology record are known but have not yet been analysed. The potential for bias in the realisation of the sample will need to be assessed once all the data are processed and linked with the CRVS data, and it is possible that a weighted analysis will need to be undertaken.

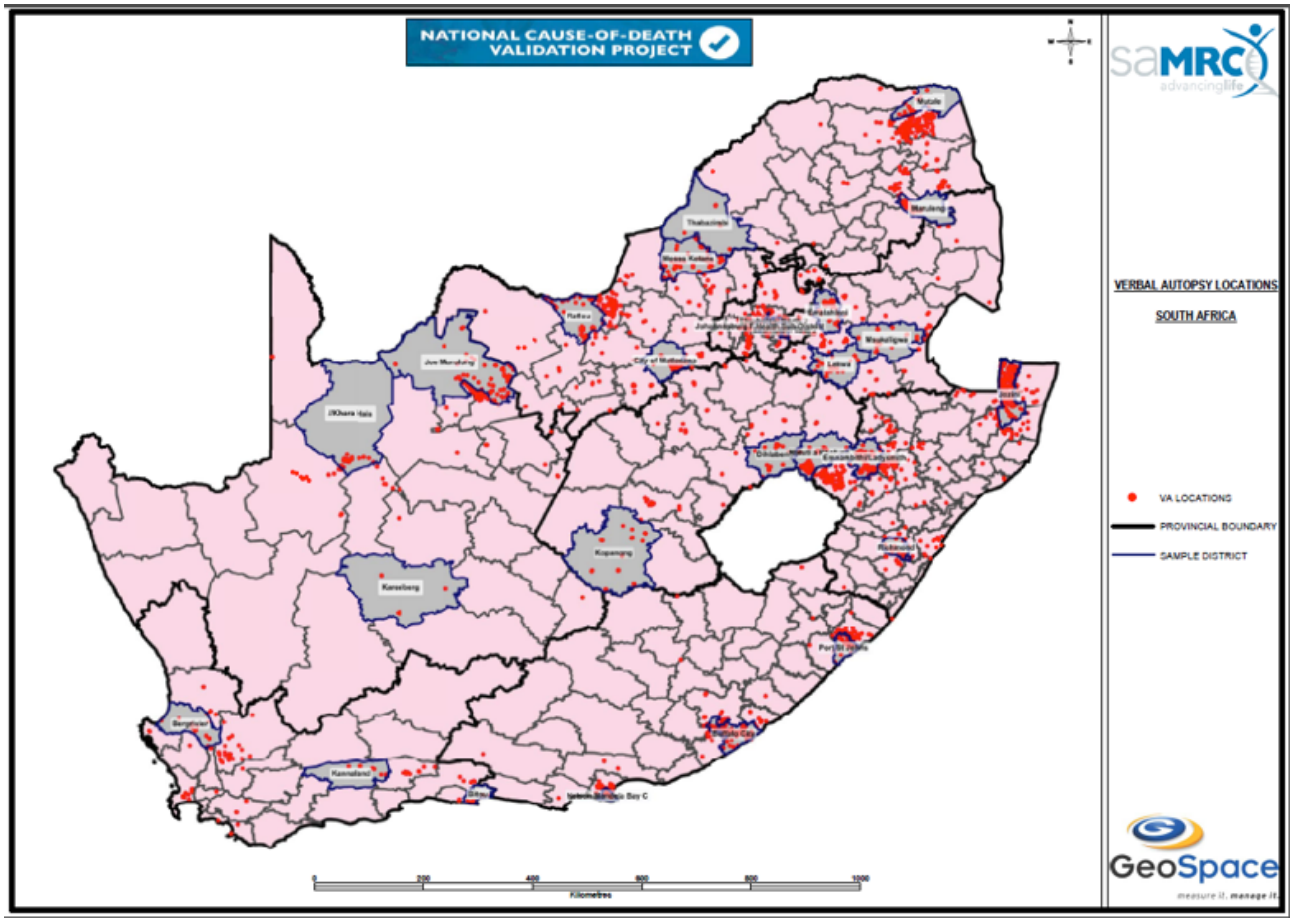


Figure 6: Geographic location of verbal autopsy interviews, SA NCOD Validation Project 2017/18.

The provincial distribution of the target sample of registered deaths is shown in Table 6 alongside the distribution of the deaths included in the study. It should be noted that the sample was not selected proportionally by province. The sample was selected to ensure that all socio-economic strata in each province were represented. Relative to the proportions of registered deaths that occurred within the study areas, deaths from Gauteng and Eastern Cape provinces are under-represented while the other provinces are over-represented, as was noted from Table 4.

Table 6: Provincial distribution of registered deaths and deaths for which data were collected, SA NCOD Validation Project 2017/18.

Province	Registered deaths (Department of Home Affairs)		Deaths included in study	
	Number of deaths	% distribution	Number of deaths	% distribution
Eastern Cape	12,960	35	6,147	23
Free State	3,442	9	2,313	9
Gauteng	9,371	25	4,262	16
KwaZulu Natal	2,240	6	2,214	8
Limpopo	635	2	1,918	7
Mpumalanga	3,114	8	2,336	9
Northern Cape	1,131	3	1,667	6
North West	3,570	10	3,449	13
Western Cape	513	1	2,208	8
Total	**36,976**	**100**	**26,514**	**100**

4.2 Verbal autopsy

4.2.1 Response rate

After removing 31 duplicate records, consent had been obtained from a total of 6,328 next of kin during the recruitment phase. Neither the next of kin nor their dwelling could be located for 560 decedents, accounting for 8.9% of the total sample. Out of the 5,768 dwellings that were located, a total of 5,387 verbal autopsies were completed (93.4%). The 380 households that refused accounted for 6.6% of the those who were approached. The overall response rate achieved was 85.2%.

Limited data are available about the characteristics of the decedents whose next of kin originally consented to be contacted but who were unable to be found or who refused. Figure 6 demonstrates the age group distribution of the 6,328 decedents whose next of kin originally consented and then responded, refused, or could not be located. The age groups were similar for the decedents for whom verbal autopsy interviews were conducted, the refusals, and the not-located. The verbal autopsy interviews had 9% more male decedents than female; the medical records had 4% more male decedents than female and the forensic records had 331% more male decedents than female. However, unlike the responders, the sex was unknown for 7.9% of the refusals and 16.1% of those not located, respectively.

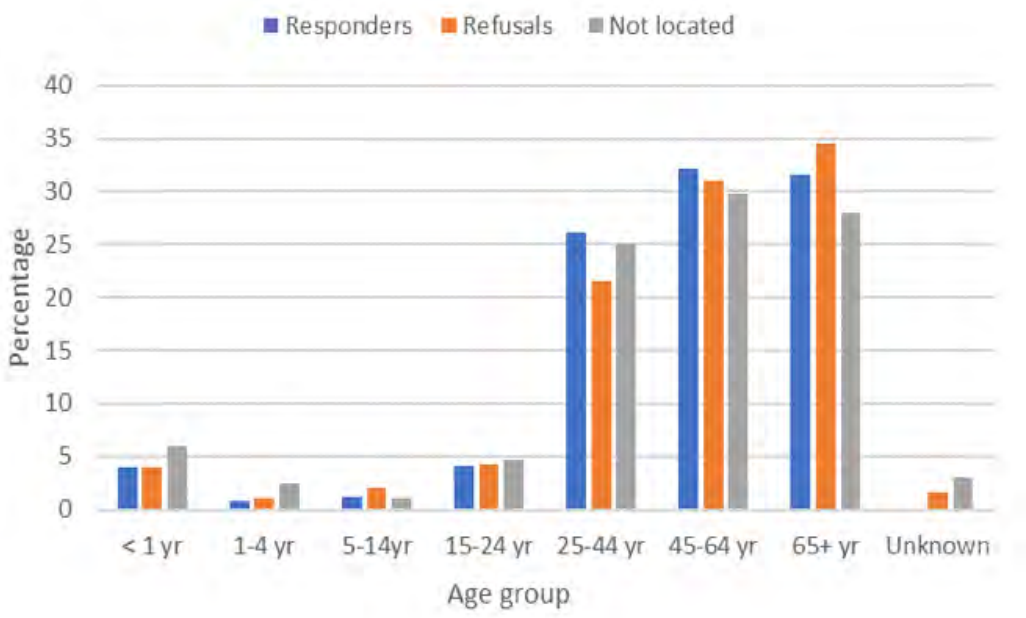


Figure 7: Age group distribution of the decedents whose next of kin responded, refused and could not be located (N=6,328), SA NCOD Validation Project 2017/18.

Once the dwelling was located, the most common reason for refusal was because the next of kin was too upset to participate (n=180, 3.1%) and a further 101 next of kin indicated that they were not interested in participating (1.7%). As shown in Table 7, there were 5 next of kin who indicated that they were concerned about legal issues (0.1%).

Table 7: Response category and reason for refusal to participate in verbal autopsy interview (N=5,768), SA NCOD Validation Project 2017/18.

Response category and reason for refusal	Number	%
Dwelling located	5,768	100.0
Verbal autopsy conducted	5,387	93.4
Refusal	380	6.6
Respondent too emotional about the death of their loved one to take part	180	3.1
Respondent had no interest in taking part in such a study	101	1.8
Respondent refused, either telephonically or face to face	47	0.8
Respondent suspicious of all surveys, citing fear of fraudulent activities and suspicion of how fieldworkers were able to obtain their personal details	25	0.4
Respondent cited a lack of time to participate	15	0.3
Information regarding the next of kin/informant incorrect, and the person could either not be contacted at all, or when contacted, did not know the respondent	6	0.1
Respondent indicated legal issues, typically in the midst of dispute with the facility where the decedent died	5	0.1

The time to complete a questionnaire varied according to the demographic of the decedent. Interviews regarding adult

male decedents took the least time, taking about 30-45 minutes to complete to questions. Adult females and children took about 45-60 minutes for the questions. However, including the time taken to make the initial contact, share information about the project and obtain consent, the overall interview could take up to 3 hours, particularly if documents like a death certificate or a road-to-health card needed to be copied.

The time interval between date of death and date of interview ranged from 3.6 months to 18.4 months with an average of 9.4 months and standard deviation of 2.7 months. For three cases, we were unable to calculate the time between death and interview because the actual day and month of death were missing from those records. A total of 838 (15.6%) verbal autopsy interviews were conducted more than 12 months after the death of the decedent. Table 8 shows the year of death reported in the verbal autopsy interviews. A higher proportion of deaths with completed verbal autopsies occurred in 2018 (68.3%).

Table 8: Year of death reported in verbal autopsy interviews (N=5,387), SA NCOD Validation Project 2017/18.

Year of death	Number	%
2017	1,671	31.0
2018	3,716	69.0
Total	**5,387**	**100.0**

4.2.2 Socio-demographic characteristics

The sex distribution of the is shown in Table 9 and the age and sex profile of the decedents with verbal autopsy interviews are shown in Figure 8 and Table 10, indicating a reasonable representation of the deaths experienced in South Africa. According to the Stats SA report of deaths in 2016, 52,1% of the registered deaths were male and 47.4% females which was very similar to the sex distribution of the sample (Table 9).

Table 9: Sex distribution of sample with verbal autopsy interviews (N=5,387), SA NCOD Validation Project 2017/18.

	Number	%	95% CI
Female	2,579	47.9	46.5-49.2
Male	2,808	52.1	50.8-53.5
Total	**5,387**	**100.0**	

The median age of death reported by Stats SA had increased from 42.8 in 2005 to 54.6 years by 2016;[55] 52.7 years for males and 62.0 years for females. The median age at death derived from the sample of verbal autopsies was 53.7 years for persons, 51.8 year for males and 56.4 years for females. While the overall median and the median of the males were similar to the Stats SA values for 2016, the female sample of decedents in this study had a notably younger median than the registered female deaths. The interquartile values from the sample were 36.2 for persons, 35.1 for males and 37.7 for females.

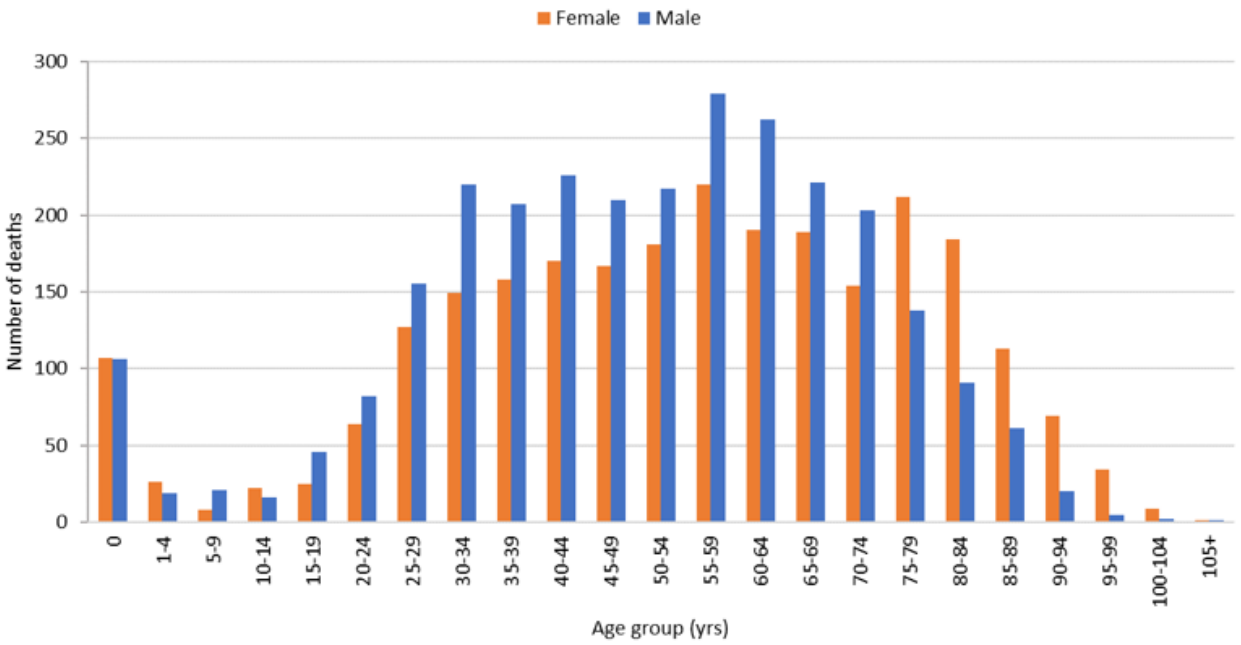


Figure 8: Age distribution of sample with verbal autopsy interviews by sex (N=5,387), SA NCOD Validation Project 2017/18.

Table 10: Age group distribution of sample with verbal autopsy interviews by sex (N=5,387), SA NCOD Validation Project 2017/18.

Age group	Male	%	Female	%	Total	%
Stillbirths	17	0.5	12	0.5	29	0.6
Early neonatal (0-6 days)	23	0.8	21	0.8	44	0.8
Late neonatal (7-27 days)	12	0.4	17	0.7	29	0.5
Post-neonatal deaths (1-11 months)	55	2.0	56	2.2	111	2.1
Child (1-4 years)	20	0.7	28	1.1	48	0.9
Older child (5-14 years)	37	1.3	30	1.2	67	1.2
Adolescent and youth (15-24 years)	128	4.6	88	3.4	216	4.0
Adult (25-44 years)	808	28.8	602	23.3	1410	26.2
Adult (45-64 years)	964	34.3	762	29.5	1726	32.0
Older adults (65+ years)	744	26.5	963	37.3	1707	31.7
Total	**2,808**	**100.0**	**2,579**	**100.0**	**5,387**	**100.0**

The population group distribution of the sample with verbal autopsy interviews indicated an over-representation of black African and coloured deaths, and an under-representation of the other population groups (Table 11). In 2016, 70.0% of all registered deaths in the country were Black African while 9% were white, 7.1% coloured, 1.9% were from the Indian/Asian population group, and 11.9% were reported as "other" and "unknown or unspecified population group."[55]

Table 11: Population group distribution of sample with verbal autopsy interviews (N=5,387), SA NCOD Validation Project 2017/18.

Population group	Number	%	95% CI
Black African	4,765	88.5	87.6-89. 2
Coloured	487	9.0	8.4-9.8
Indian/Asian	19	0.4	0.2-0.6
White	116	2.2	1.8-2.6
Total	**5,387**	**100.0**	

CI = confidence interval

4.2.3 Place and province of death

Table 12 shows the place of death of the sample. The majority of the deaths occurred in hospital (52.3%). This is higher than the 43.2% observed in the 2016 registered deaths reported by Stats SA. However, registered deaths in 2016 Stats SA have a high proportion of missing information about the place of death (23.2%) compared to this study (0.2%).[55]

Table 12: Place of death of sample with verbal autopsy interviews (N=5,387), SA NCOD Validation Project 2017/18.

Place of death	Number	%	95% CI
Don't know/missing	13	0.1	0.07-0.30
Home	1,995	37.1	35.8-38.4
Hospital	2,817	52.3	51.0-53.7
On route to hospital	122	2.3	1.9-2.7
Other health facility	110	2.0	1.7-2.5
Other	330	6.1	5.5-6.8
Total	**5,387**	**100.0**	

CI = confidence interval

The provincial distribution of the sample with verbal autopsy reflects a combination of the sample that was drawn, the success of the recruitment in each area, and the response rate for the verbal autopsy. Table 12 shows the provincial breakdown of the sample as well as the provincial breakdown of the Stats SA 2016 deaths (including stillbirths but excluding deaths with unknown province of death and deaths that occurred outside of South Africa). It can be seen from Table 13 that all provinces were represented in the sample, but the proportions of deaths within the sample were not similar to the proportion of total deaths in South Africa experienced in each province, reflecting that the response rate was not even across all provinces. The proportion of the sample for Gauteng was much lower than anticipated. While the recruited funeral practitioners expressed willingness to support the project, the number next of kin recruited to the study was extremely low. Recruitment in Western Cape was also challenging as it was difficult to identify the funeral practitioners who serviced the sampled areas which were all in rural areas.

Table 13: Province of death of sample with verbal autopsy interviews (N=5,387) compared with Stats SA 2016, SA NCOD Validation Project 2017/18.

Province	Number	% of total deaths in sample	Stats SA 2016	% of total deaths in South Africa
Eastern Cape	575	10.7	66,607	14.2
Free State	655	12.2	32,746	7.0
Gauteng	463	8.6	100,390	21.4
KwaZulu Natal	891	16.5	87,414	18.7
Limpopo	377	7.0	46,880	10.0
Mpumalanga	610	11.3	34,170	7.3
Northern Cape	506	9.4	14,229	3.0
Northwest	1,108	20.6	36,454	7.8
Western Cape	202	3.7	49,378	10.5
Total	**5,387**	**100.0**	**468,268**	**100.0**

* *Includes 11,960 registered stillbirths & excludes 252 deaths outside South Africa and 52 with unknown province*

4.2.4 Cause of death profile based on InterVA-5

Out of the 5,387 verbal autopsy interviews, there were 10 stillbirths identified by InterVA-5 based on the questionnaires without reference to information in the narrative. Six were fresh stillbirths and four were macerated stillbirths. All ten have been excluded from further analysis (N=5,377).

The cause of death profile based on the most probable cause identified by InterVA-5 for the 5,377 cases is shown in Figure 9. A cause could not be identified for 254 of the deaths (4.7%). HIV/AIDS was the most common cause and accounted for 20.6% of the all the deaths and pulmonary TB for 8.7%. Stroke accounted for 8.8% of deaths and other unspecified cardiac conditions for 8.5%. All injuries combined accounted for 12.6% of the deaths.

The cause profiles differ between males and females (Figure 10). Males had a higher proportion of pulmonary TB deaths (11.3%) compared to females (5.9%) while females have a higher proportion of HIV/AIDS related deaths (23.1% vs 18.3%). Females had higher proportion of stroke (10.5%) than males (7.3%) as well as other unspecified cardiac diseases (11.1% vs 6.0%). Injuries accounted for 17.9% of male deaths and 6.8% of female deaths.

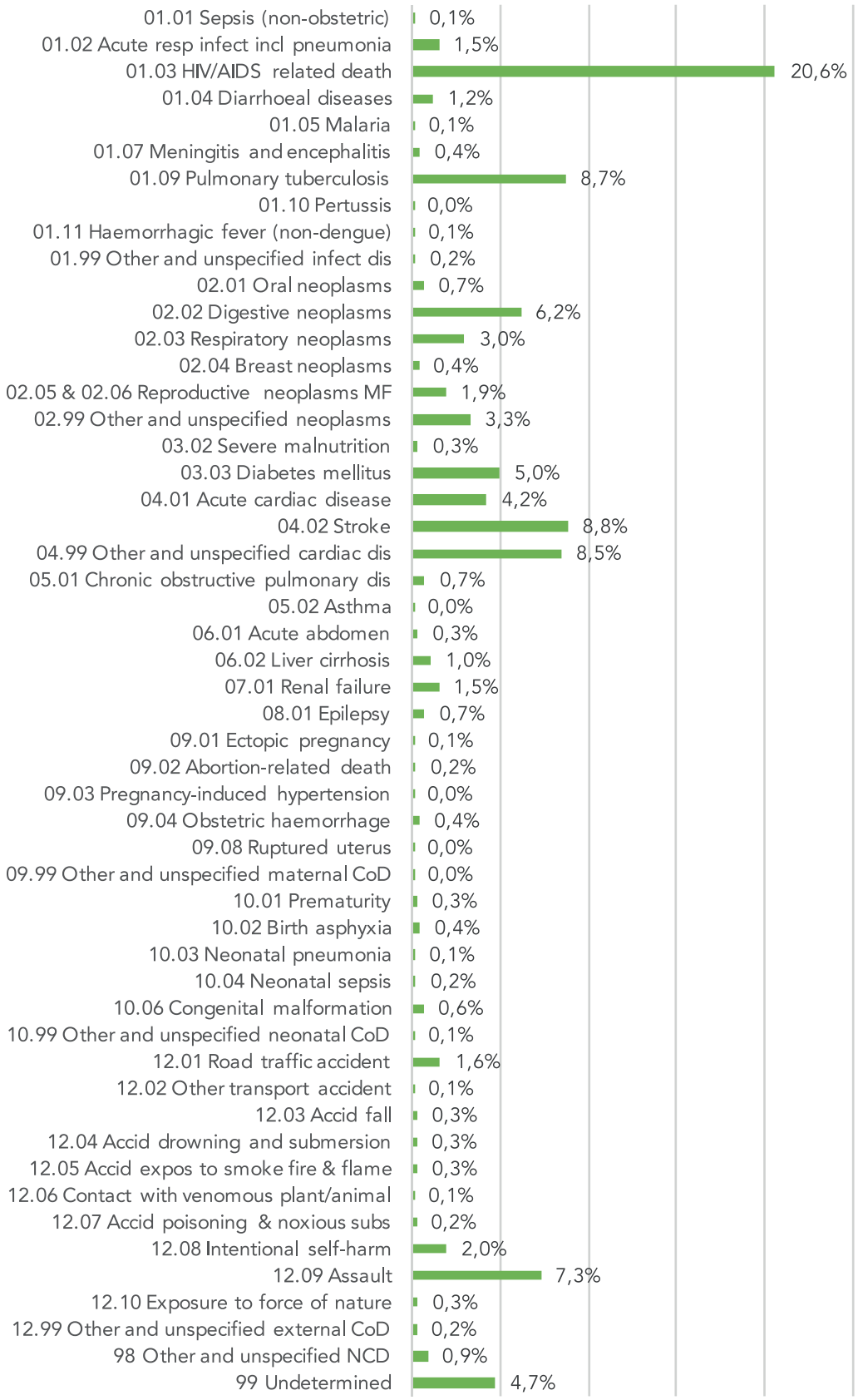


Figure 9: Cause of death from verbal autopsy interviews based on InterVA-5 (N=5,377), SA NCOD Validation Project 2017/18.

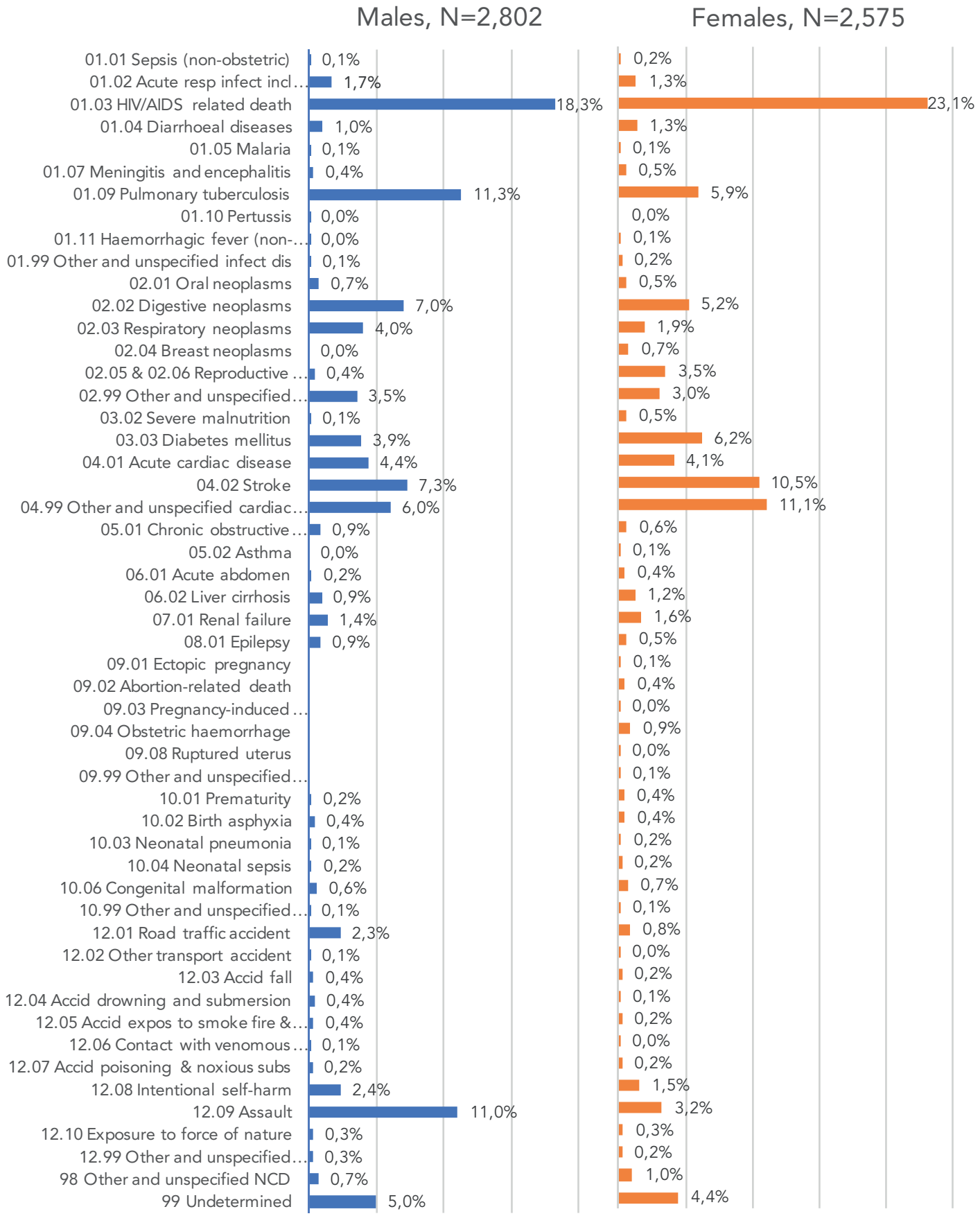
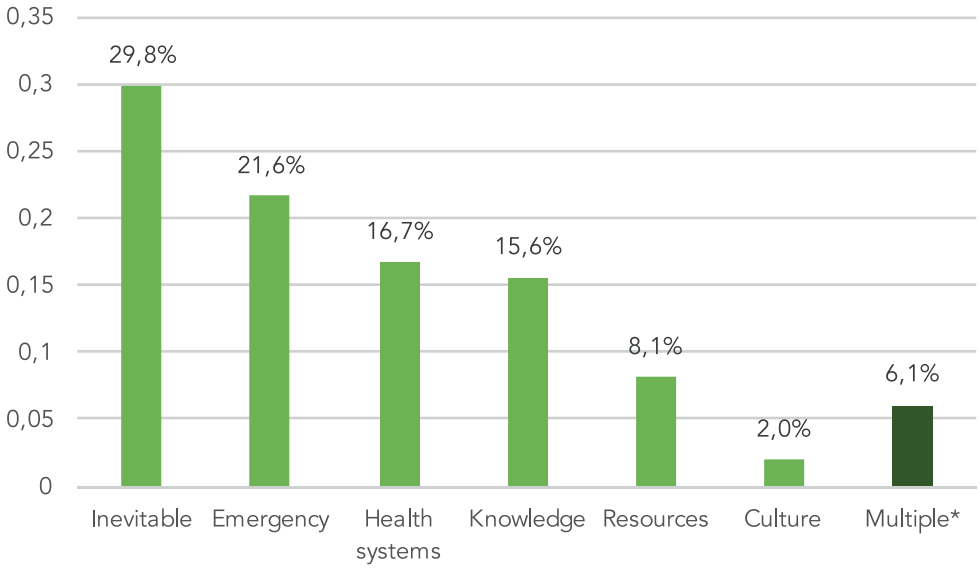


Figure 10: Male and female cause of death from verbal autopsy interviews based on InterVA-5 (N=5,377) SA NCOD Validation Project 2017/18.

Figure 11 shows that about half of the deaths were considered inevitable or emergencies which probably precluded life-saving actions. However, 16.7% were related to problems getting health care despite accessing health facilities (e.g. related to admissions, treatments, and medications) and 8.1% were related to hindered access to health care as a result of an inability to mobilise and use resources (e.g. material, transport, financial). A further 15.5% were related to poor health seeking behaviour caused by a lack of recognition or awareness of serious disease (e.g. symptoms or severity). Cultural practices or beliefs influenced health seeking behaviours and the pathway to death in a relatively low proportion of cases (2.0%). Since 6.1% of the deaths were categorised into multiple categories, it is likely that the proportions for each of the categories were higher. There was a noticeably lower proportion of emergency conditions experienced by females (16.5%) vs males (26.3%), as can be seen from Figure 12.



Multiple indicates 2 or more circumstances*

Figure 11: Circumstances of mortality categories (COMCATs) based on verbal autopsy interviews (N=5,377), SA NCOD Validation Project 2017/18.

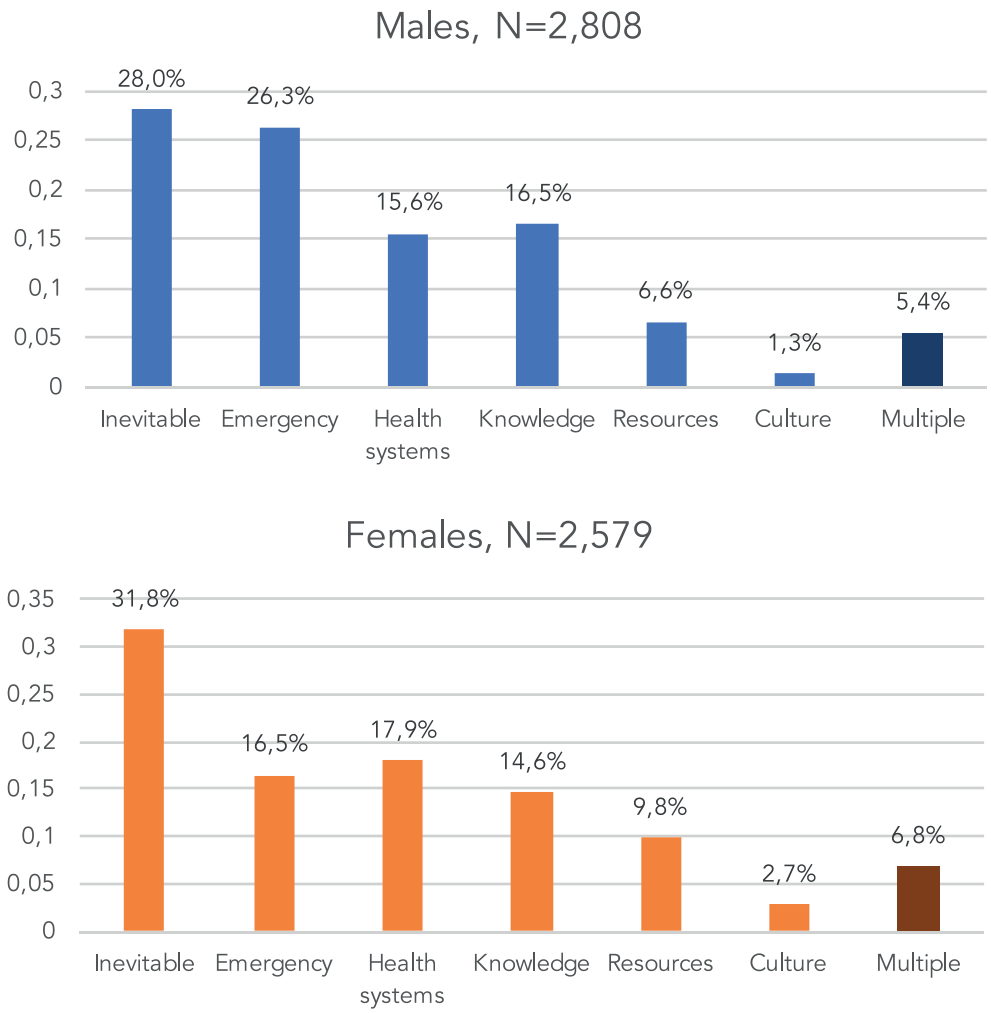


Figure 12: Male and female circumstances of mortality categories (COMCATs) based on verbal autopsy interviews (N=5,377), SA NCOD Validation Project 2017/18.

4.2.5 Cause of death profile based on doctor review

A total of 5,429 VA records and narratives, including 24 duplicate records (same USID number), were reviewed independently by two trained doctors. Out of the total of 5,417 individual records 3,400 (62.8%) cases, there was agreement on the causal sequence of death between the independent reviews. In 2,017 (37.2 %) cases, the two doctors arrived at a consensus review and 556 (10.3%) cases had to be referred to the panel for review and assignment of the causal sequence.

Further checking of the 5417 deaths revealed that 30 were duplicate records which had been assigned different USIDs. From the remaining 5,387 verbal autopsy interviews the doctors identified 29 stillbirths. The stillbirths are presented separately. The quality of the underlying cause information of the remaining 5,358 deaths was assessed using the criteria developed by an expert group convened by the Bloomberg Philanthropies Data for Health Initiative and the Civil Registration and Vital Statistics Improvement project of the University of Melbourne in 2017.[100] The high quality of cause of death certification information provided by the study doctors can be observed from Figure 13. More than two thirds of the deaths were coded to usable codes. Although 16.6% of the causes are considered to have insufficient specification with an ICD chapter, this should be anticipated as the information was derived from a verbal autopsy interview with a lay person. The analysis is presented by sex in Figure 14. The use of intermediate causes of death as the underlying cause occurred in a slightly higher proportion for females and the resultant proportion of usable codes was slightly lower for females (65.8%) compared to males (70.5%).

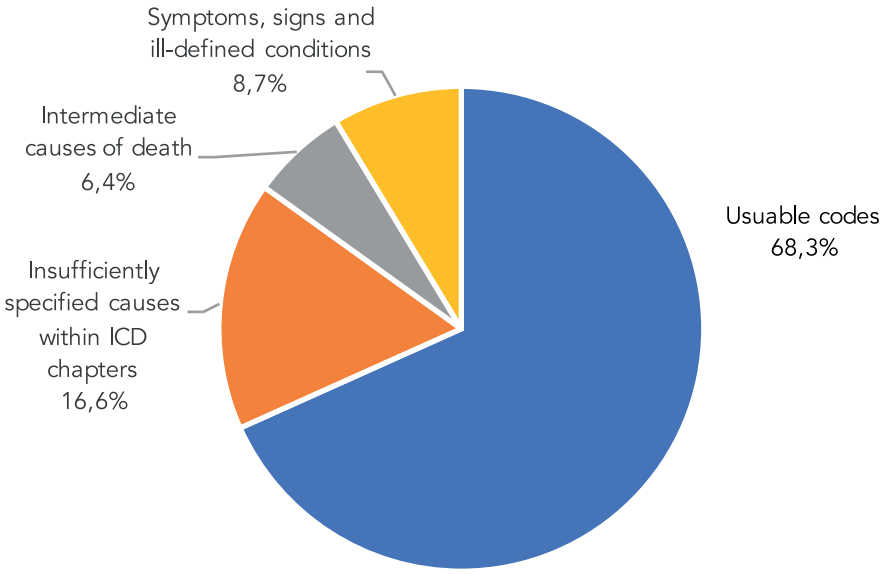


Figure 13: Assessment of the underlying cause of death data from doctor reviewed verbal autopsies (N=5,358), SA NCOD Validation Project 2017/18.

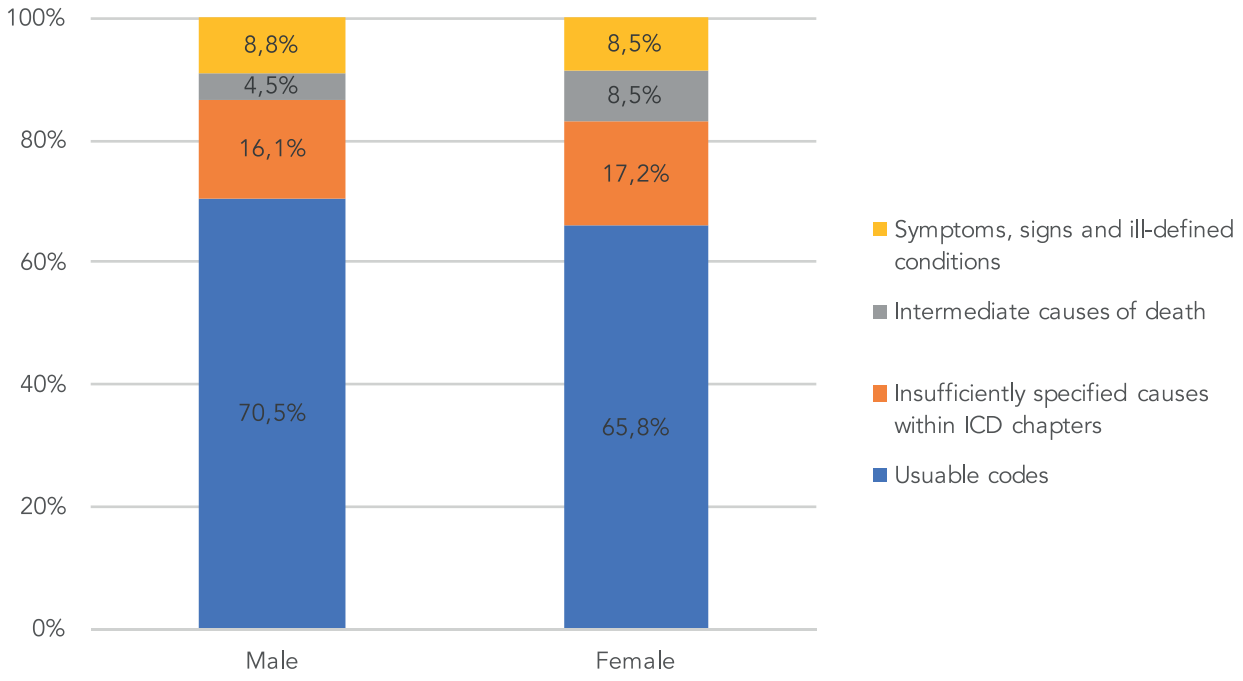


Figure 14: Assessment of underlying cause of death data by sex (N=5,358), SA NCOD Validation Project 2017/18.

The overall cause of death profile by ICD chapter is shown in Figure 15. The chapter for infectious and parasitic diseases accounted for 31.4% of all the deaths followed by the circulatory chapter (19.8%). External causes accounted for 12.8% and ill-defined causes and symptoms for 8.7%. From Figure 16 it can be seen that the profile for males and females differed, with females having higher proportions of deaths due to circulatory conditions (26.4%) compared to that for males (14.9%). As expected, the proportion of external causes was much higher for males (20.0%) than for females (6.8%). In addition, infectious and parasitic conditions were more common among male deaths (36.3%) than among female deaths (29.2%).

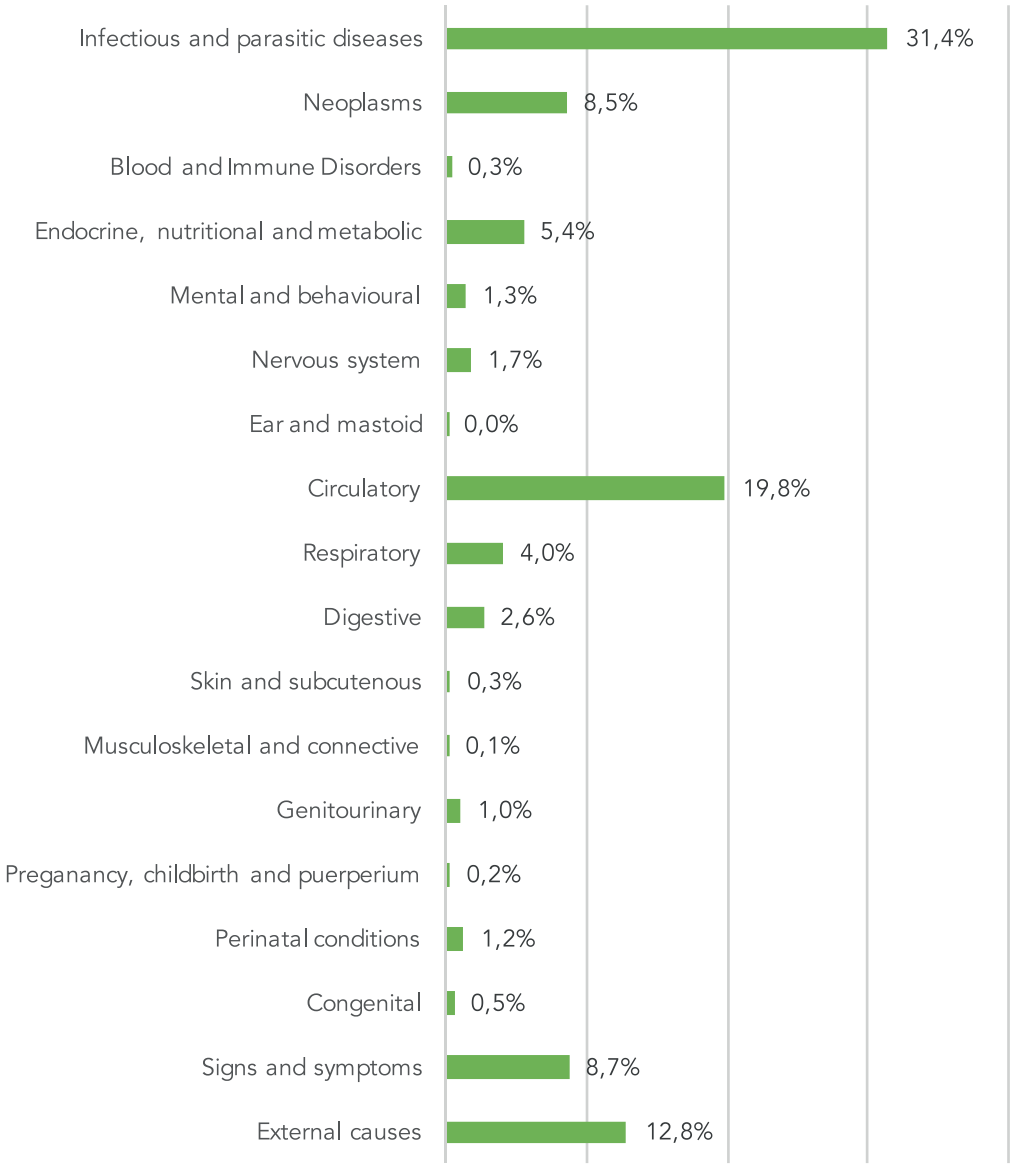


Figure 15: Cause of death by ICD chapter based on doctor review of verbal autopsy interviews (N=5,358), SA NCOD Validation Project 2017/18.

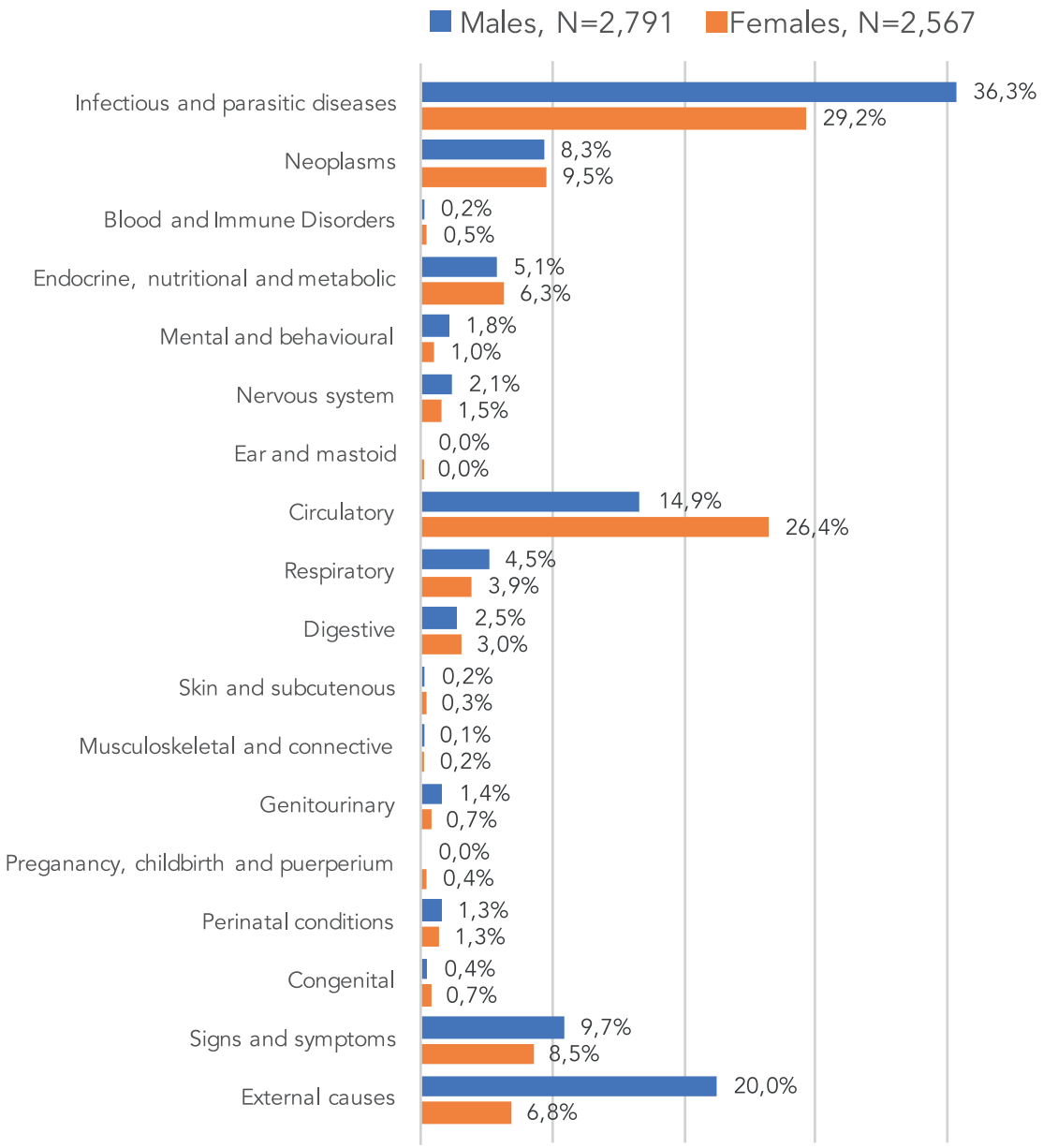


Figure 16: Male and Female cause of death by ICD chapter based on doctor review of verbal autopsy interviews (N=5,358), SA NCOD Validation Project 2017/18.

HIV was the most commonly identified cause, accounting for 22.8% of all deaths. HIV and TB accounted for 29.9% of the sample of deaths (N=1,601). The specific causes are shown in Figure 17, indicating the combination of HIV resulting in TB was the most common cause accounting for 37.7% of the HIV and TB deaths. Overall, 61.5% of the TB deaths were related to HIV. The HIV related causes are shown in Figure 18 which highlights that nearly half of the HIV deaths occurred with TB (49.3%). The proportion of HIV and TB deaths that had mention of defaulting on treatment is shown in Table 14. Out of the 1,223 deaths from HIV/AIDS, 10.2% had mention of treatment default while 4.0% of the TB deaths had mention of treatment.

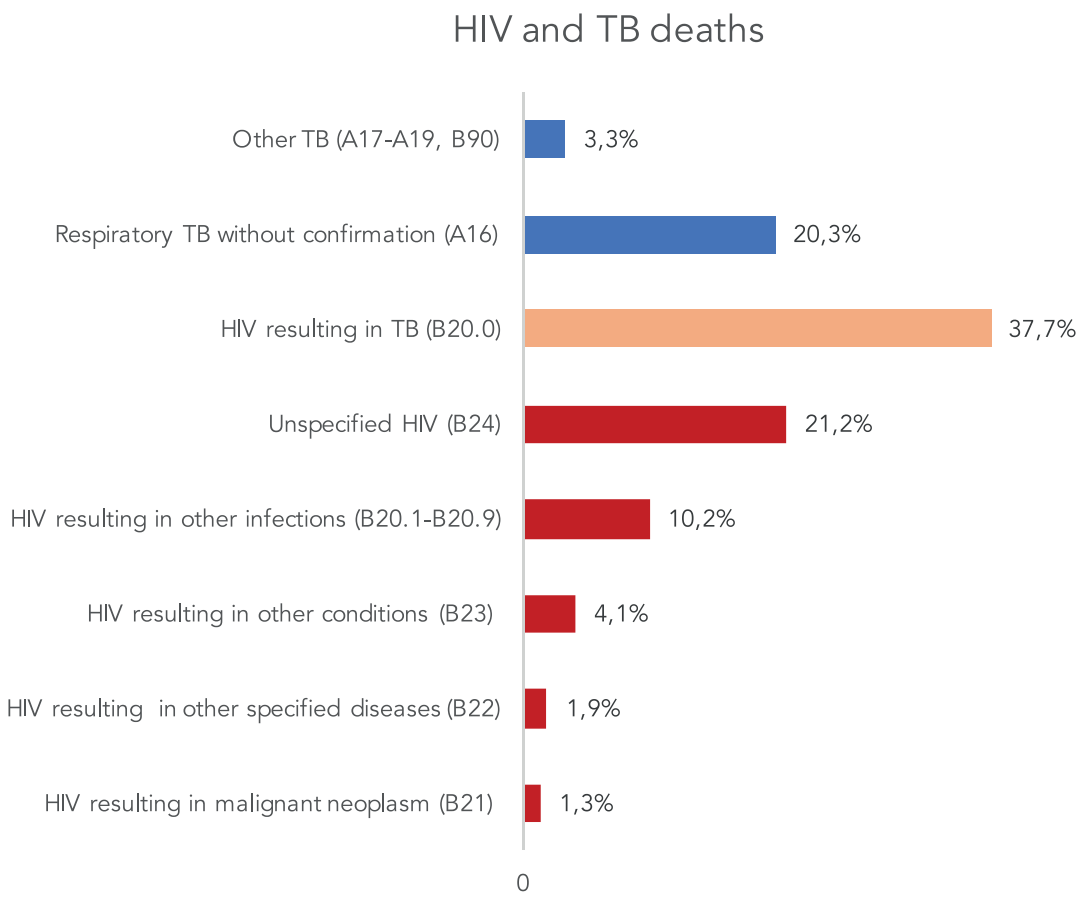


Figure 17: Distribution of HIV and TB related deaths by ICD codes based on doctor review of verbal autopsy interviews (N=1,601), SA NCOD Validation Project 2017/18.

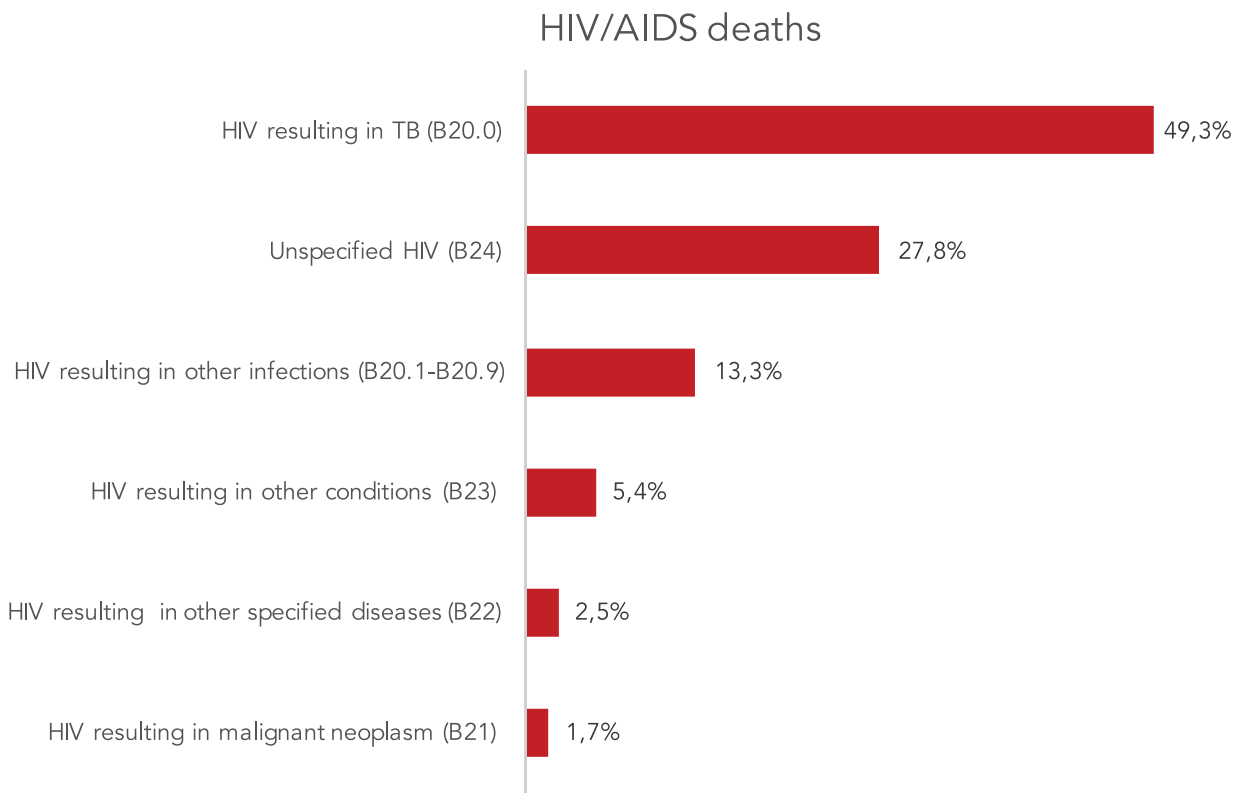


Figure 18: Distribution of HIV related deaths by ICD codes based on doctor review of verbal autopsy interviews (N=1,223), SA NCOD Validation Project 2017/18.

Table 14: Number and percentage of HIV and TB decedents with mention of defaulting treatment in brief medical history based on the verbal autopsy, SA NCOD Validation Project 2017/18.

Cause of death	Defaulted treatment	%	Total
TB (A16-A19; B90) deaths	15	4.0	378
HIV (B20-B24) deaths	125	10.2	1,223
HIV and TB (A16-A19; B20-B24; B90) deaths	140	8.7	1,601

A breakdown of the 688 injury-related causes of death is shown in Figure 19. Homicides were the leading single cause of death, accounting for 31.0% of all injury deaths, followed by transport injuries (20.5%) when using the national burden of disease categorisation of injuries. This was followed by suicides which accounted for 9.2% of the injury related deaths. The profiles for males and females were similar for the leading causes with differences in the proportions for selected causes (Figure 20). For example, surgical and medical misadventure accounted for 4.6% of the female injury deaths compared with 0.8% of male deaths. However, these are based on small numbers and should not be overinterpreted (8 female and 4 males).

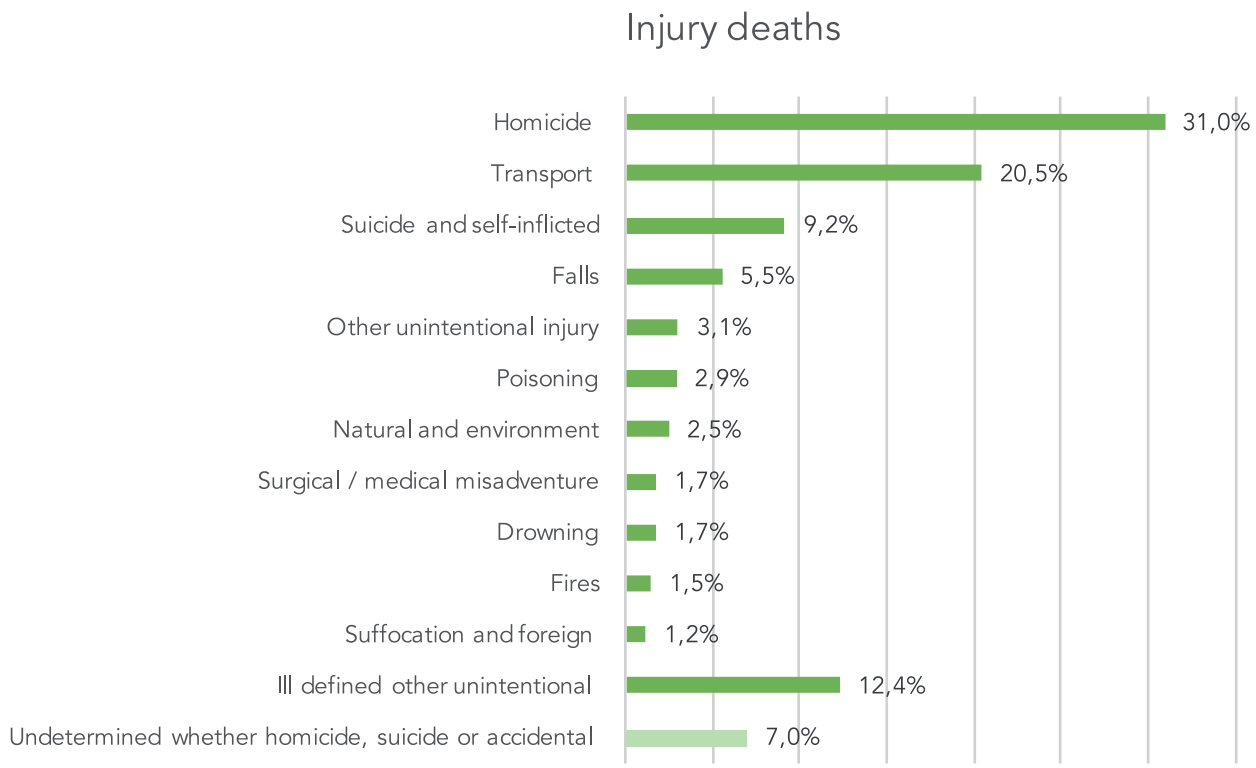
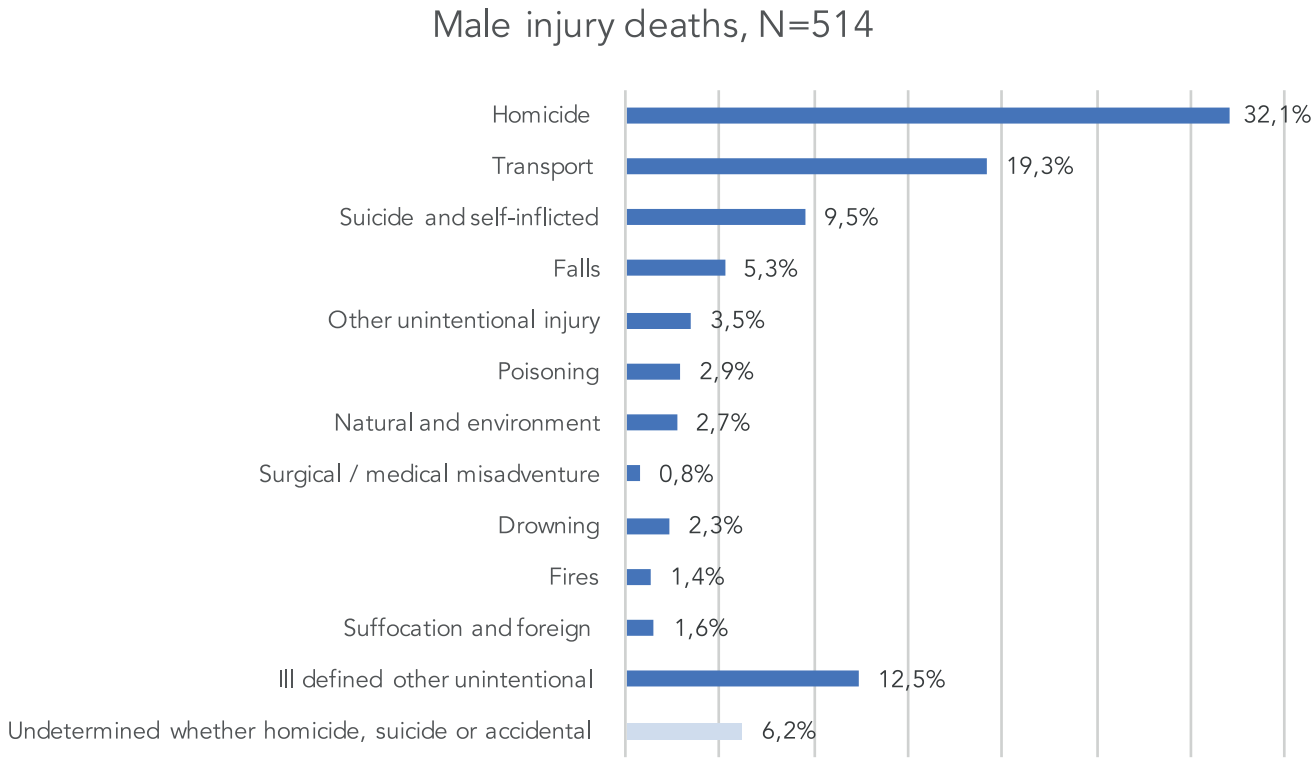


Figure 19: Injury-related causes of death based on doctor review of verbal autopsy interviews (N=688), SA NCOD Validation Project 2017/18.



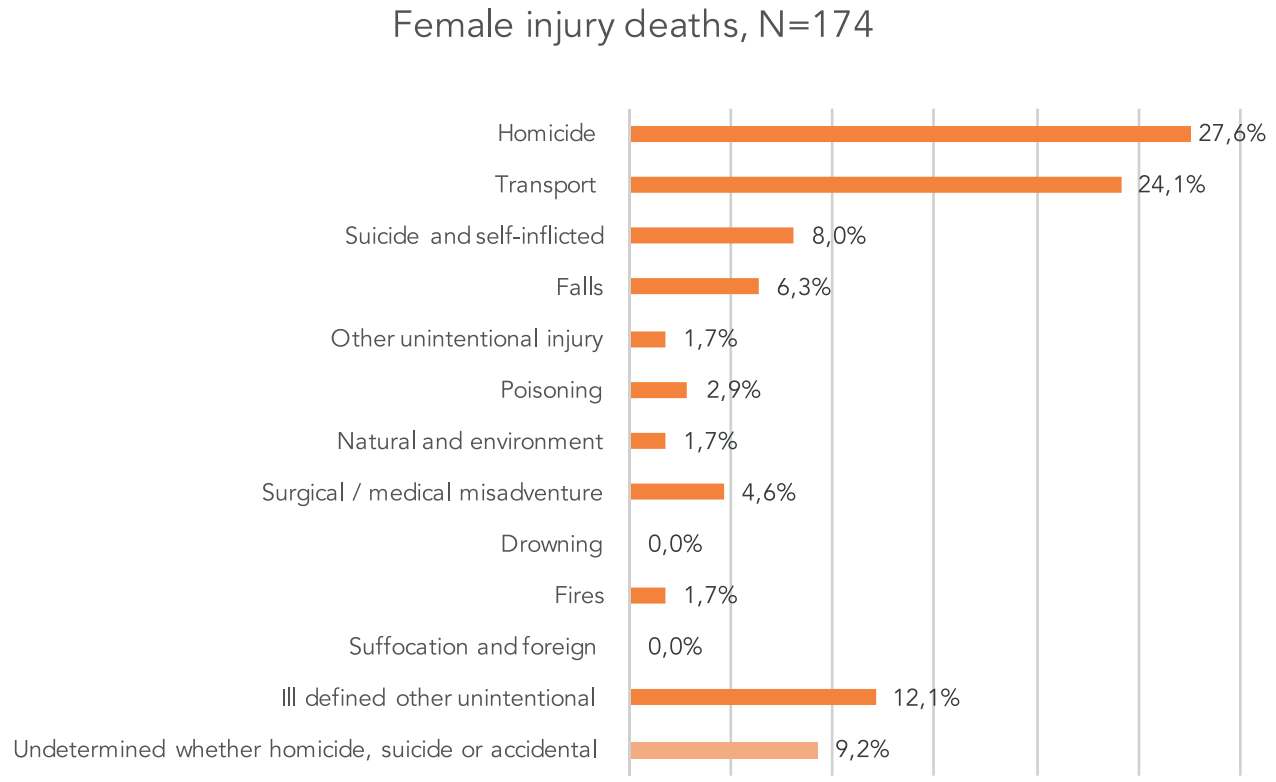


Figure 20: Male and female injury-related causes of death based on doctor review verbal autopsy interviews, SA NCOD Validation Project 2017/18.

An investigation into the 29 stillbirths found that the majority of the verbal autopsies did not provide adequate information about the underlying cause of death. It can be seen in Figure 21, a high proportion are ill-defined causes (31.0%) and only 41.4% of the causes were usable. The breakdown of the stillbirth causes of death is shown in Table 15. However, the data are very sparse and should be interpreted with caution.

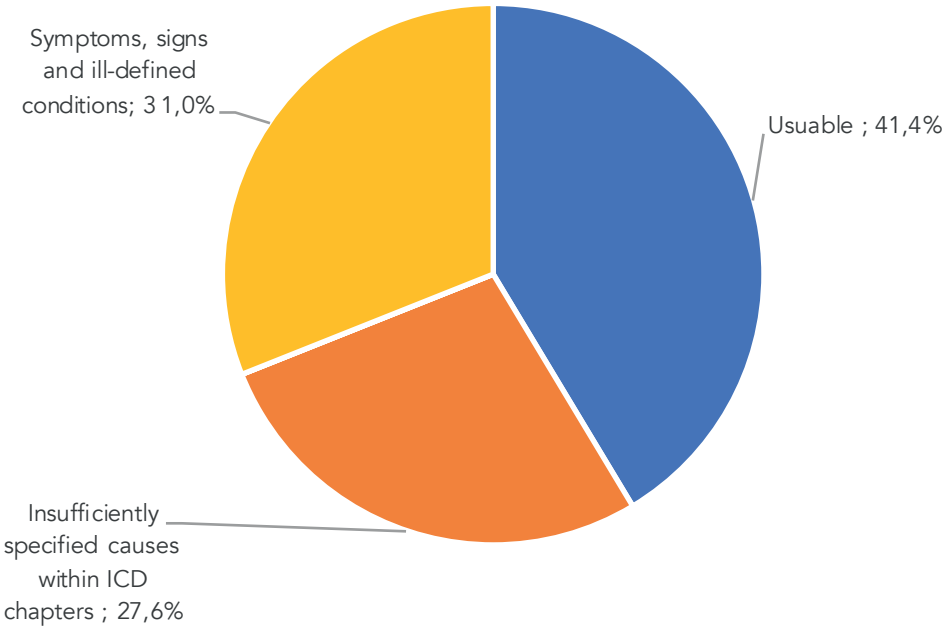


Figure 21: Assessment of the stillbirth underlying cause of death data from doctor reviewed verbal autopsies (N=29), SA NCOD Validation Project 2017/18.

Table 15: Stillbirth causes of death based on doctor reviewed verbal autopsies (N=29), SA NCOD Validation Project 2017/18.

Stillbirth cause of death	Number	%
P95 Fetal death of unspecified cause (R99)	9	31.0
P96 Other conditions originating in the perinatal period	6	20.7
P07 Disorders related to short gestation and low birth weight, not elsewhere classified	4	13.8
P02 Fetus and newborn affected by complications of placenta, cord and membranes	2	6.9
P08 Disorders related to long gestation and high birth weight	2	6.9
P20 Intrauterine hypoxia	2	6.9
P37 Other congenital infectious and parasitic diseases	2	6.9
P05 Slow fetal growth and fetal malnutrition	1	3.4
Q24 Other congenital malformations of heart	1	3.4
Total	**29**	**100.0**

4.3 Comparison of cause of death profile of study sample based on verbal autopsies (automatic coding from InterVA-5 and from doctor reviews) with 2016 Stats SA

The 2017 cause of death data has not been reported by Stats SA by December 2019. Therefore the 2016 data are presented as a comparator for the study sample. Figure 22 shows the broad cause of death profile for the verbal autopsies based on automated identification of the underlying cause using InterVA-5 as well as the doctor reviews that were coded using Iris as compared to Stats SA 2016. While all three groups had a similar proportion of injury deaths, Stats SA had a much lower proportion of HIV/AIDS and TB deaths and higher proportions of the other group 1 conditions (communicable diseases: other infections, infections, maternal, neonatal and nutritional conditions), unknown causes, and non-communicable diseases. Figure 23 shows similar differences between Stats SA and verbal autopsy results despite the strong gender patterns in the cause of death profiles.

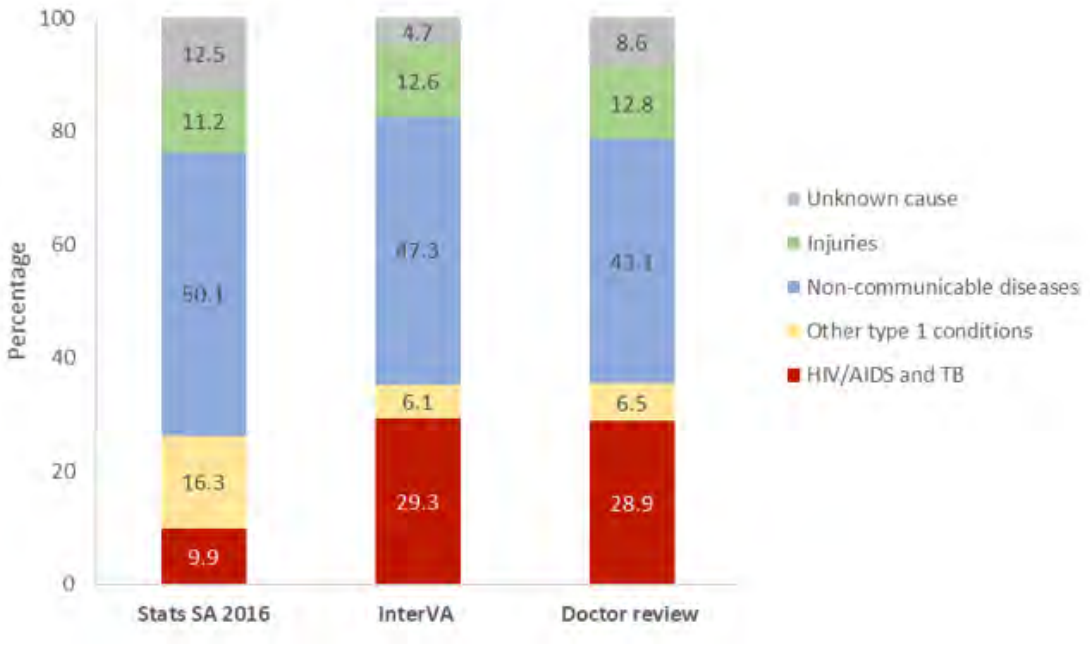


Figure 22: Broad cause group based on verbal autopsies (InterVA-5 and doctor review) compared with Stats SA 2016, SA NCOD Validation Project 2017/18.

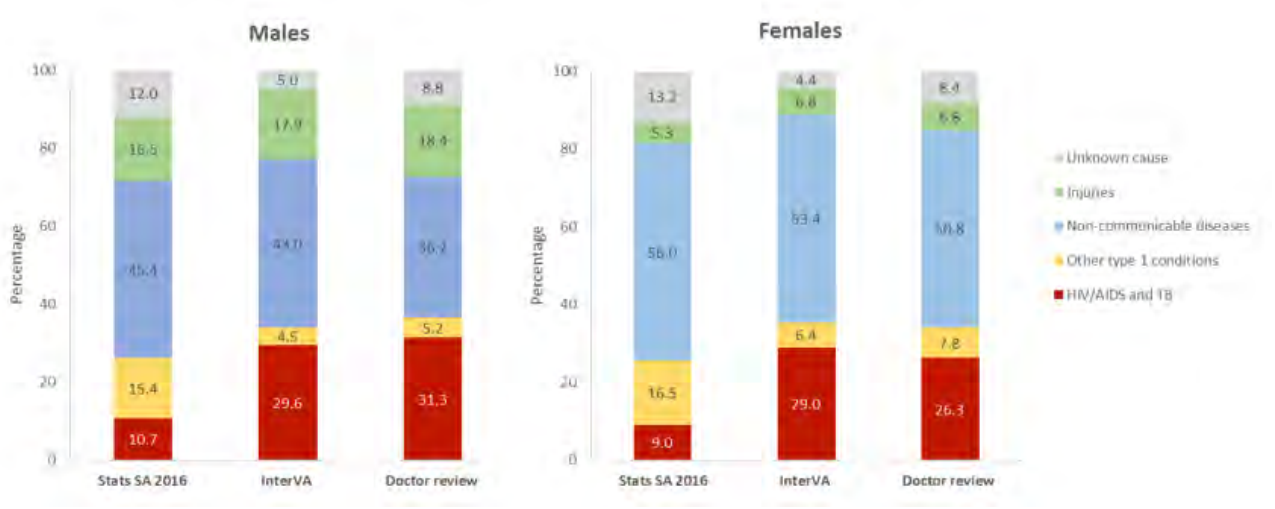


Figure 23: Broad cause group based on verbal autopsies (InterVA-5 and doctor review) compared with Stats SA 2016 by sex, SA NCOD Validation Project 2017/18.

Figure 24 shows the chapter level cause of death profiles based on Stats SA 2016 data and the proportions are compared in Table 16. Although there are some similarities between the sample of verbal autopsies and the national data for 2016, some differences are apparent at the chapter level. The proportion of deaths due to injuries were similar as well as the proportion due to circulatory conditions. However, the verbal autopsy (regardless whether automated with InterVA-5 or doctor reviewed) identified a higher proportion of infectious and parasitic disease than Stats SA (32.9% and 31.4% vs 18.2%). This may be a result of influenza and pneumonia being included in the Respiratory chapter in the Stats SA data (accounting for 9.4%) while the verbal autopsy was <1% and 4.0% for InterVA-5 and doctor reviewed respectively. The profile of the cause of death based on Stats SA 2016 data at the VA list level is shown in Figure 25 and is compared with the verbal autopsy data in Table 17. It can be seen that the verbal autopsy identified a much higher proportion of HIV related deaths (20.6% and 22.8% respectively) compared with 4.8% in the Stats SA data. In addition, slightly higher proportions of pulmonary TB deaths were identified by the verbal autopsy approach. The higher proportion of deaths due to stroke based on InterVA-5 compared with the doctor reviewed verbal autopsy (8.8% vs 5.7%) points to possible over-estimation of stroke by InterVA-5. While the total proportion of injuries are comparable, the higher proportion of assault from InterVA-5 data (7.3%) and doctor reviewed verbal autopsy (4.3%) are somewhat higher than the Stats SA data (1.7%). It is difficult to evaluate these differences which may arise from the selection in the sample or the methodology for assigning cause of death and it will be essential to undertake the planned data linkages in order to evaluate the correspondence.

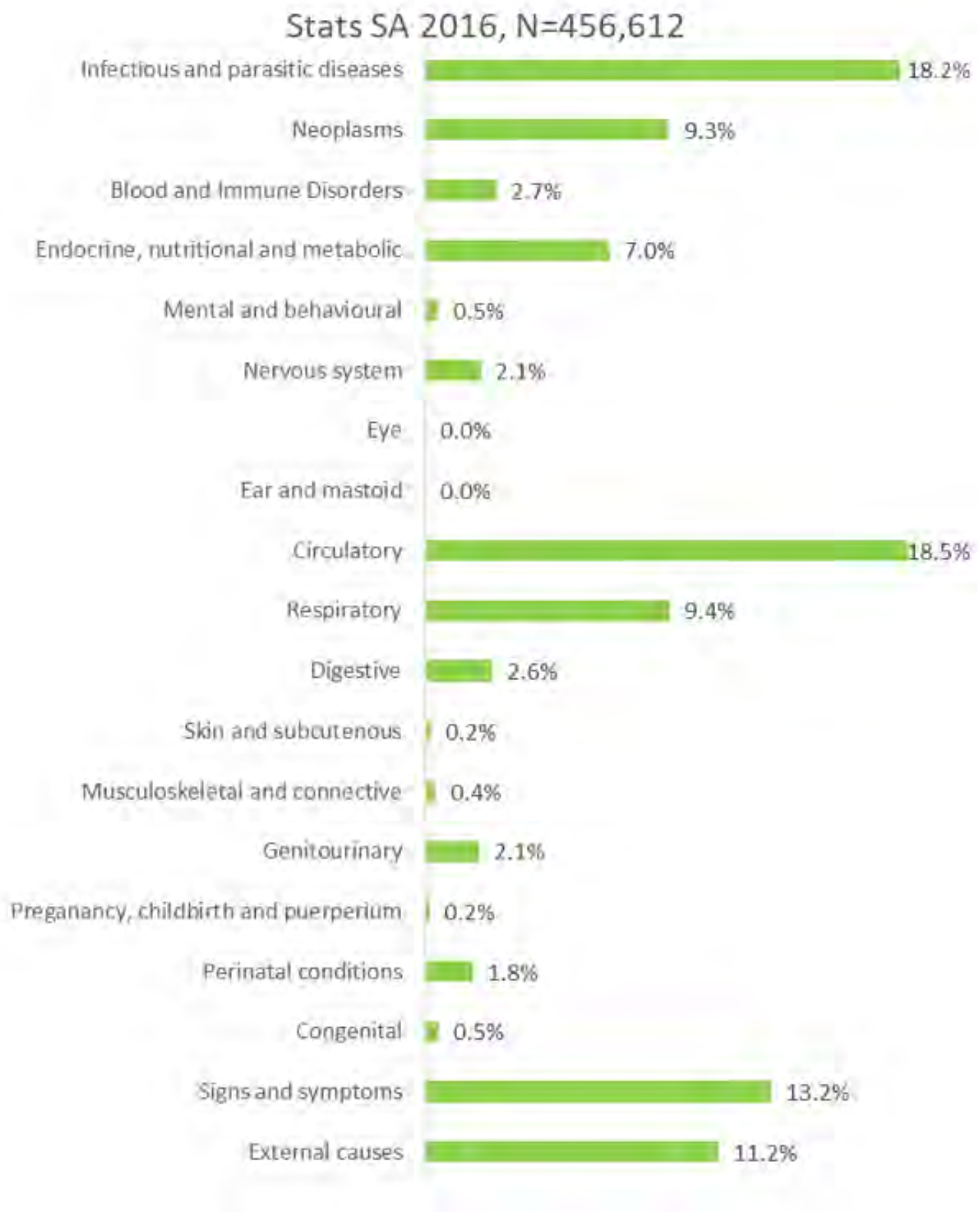


Figure 24: Cause of death by ICD chapter 2016 Stats SA data (N=456,612).

Source: Own analysis of data from Stats SA.

Table 16: Comparison of cause of death profiles based on verbal autopsies (InterVA-5 and doctor review) compared with Stats SA 2016 by ICD chapter, SA NCOD Validation Project 2017/18.

ICD chapter	Stats SA 2016	Verbal autopsy 2017/18	
		InterVA-5	Doctor reviewed and IRIS coded
N	**456,612**	**5,378**	**5,358**
Infectious and parasitic diseases	18.2%	32.9%	31.4%
Neoplasms	9.3%	15.3%	8.5%
Blood and Immune Disorders	2.7%	0.0%	0.3%
Endocrine, nutritional and metabolic	7.0%	5.3%	5.4%
Mental and behavioural	0.5%	0.0%	1.3%
Nervous system	2.1%	0.7%	1.7%
Eye	0.0%	0.0%	0.0%
Ear and mastoid	0.0%	0.0%	0.0%
Circulatory	18.5%	21.5%	19.8%
Respiratory	9.4%	0.8%	4.0%
Digestive	2.6%	1.4%	2.6%
Skin and subcutaneous	0.2%	0.0%	0.3%
Musculoskeletal and connective	0.4%	0.0%	0.1%
Genitourinary	2.1%	1.5%	1.0%
Pregnancy, childbirth and puerperium	0.2%	0.7%	0.2%
Perinatal conditions	1.8%	1.0%	1.2%
Congenital	0.5%	0.7%	0.5%
Signs and symptoms	13.2%	5.6%	8.7%
External causes	11.2%	12.6%	12.8%
Total	**100.0%**	**100.0%**	**100.0%**

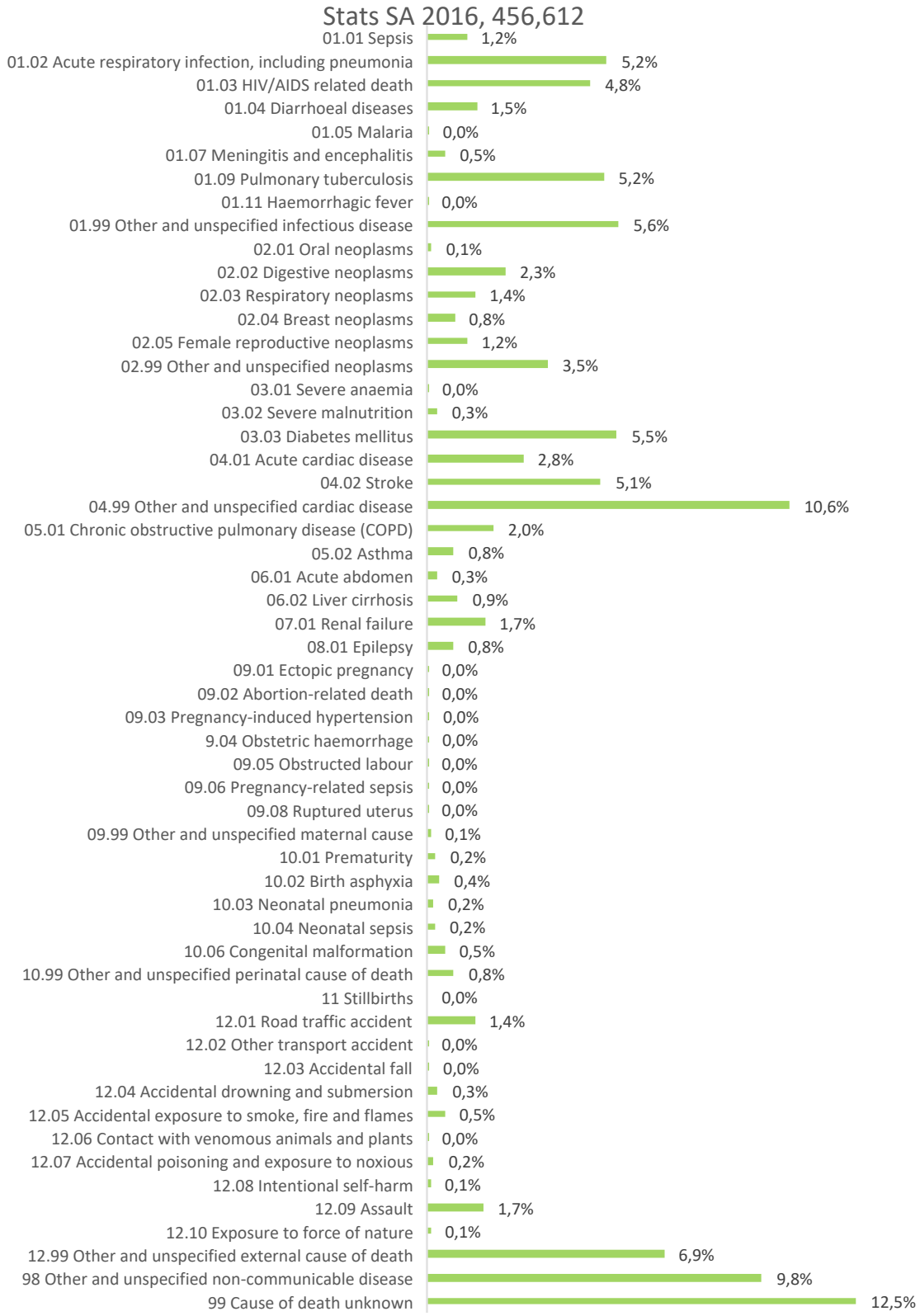


Figure 25: Cause of death by VA cause list 2016 Stats SA (N=456,612).

Source: Own analysis of data from Stats SA.

Table 17: Comparison of cause of death profiles based on verbal autopsies (InterVA-5 and doctor reviewed) compared with Stats SA 2016 by VA list, SA NCOD Validation Project 2017/18.

	Stats SA 2016			Verbal autopsy NCOD Validation 2017/18					
				InterVA-5			Doctor reviewed / IRIS coded		
VA list	Person	Male	Female	Person	Male	Female	Person	Male	Female
N	456,612	240,001	214,988	5,377	2,802	2,575	5,358	2,791	2,567
01.01 Sepsis	1.2%	1.0%	1.3%	0.1%	0.1%	0.2%	0.2%	0.1%	0.3%
01.02 Acute respiratory infection, including pneumonia	5.2%	5.2%	5.3%	1.5%	1.7%	1.3%	1.3%	1.0%	1.7%
01.03 HIV/AIDS related death	4.8%	4.6%	5.0%	20.6%	18.3%	23.1%	22.8%	23.4%	22.2%
01.04 Diarrhoeal diseases	1.5%	1.3%	1.6%	1.2%	1.0%	1.3%	1.0%	0.7%	1.2%
01.05 Malaria	0.0%	0.0%	0.0%	0.1%	0.1%	0.1%	0.2%	0.2%	0.2%
01.07 Meningitis and encephalitis	0.5%	0.5%	0.5%	0.4%	0.4%	0.5%	0.6%	0.5%	0.7%
01.09 Pulmonary tuberculosis	5.2%	6.2%	4.0%	8.7%	11.3%	5.9%	6.1%	7.9%	4.1%
01.11 Haemorrhagic fever	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
01.99 Other and unspecified infectious disease	5.6%	5.5%	5.7%	0.1%	0.0%	0.1%	1.1%	1.1%	1.2%
02.01 Oral neoplasms	0.1%	0.2%	0.1%	0.2%	0.1%	0.2%	0.0%	0.1%	0.0%
02.02 Digestive neoplasms	2.3%	2.5%	2.1%	0.7%	0.7%	0.5%	1.7%	1.6%	1.8%
02.03 Respiratory neoplasms	1.4%	1.8%	1.0%	6.2%	7.0%	5.2%	0.8%	0.9%	0.6%
02.04 Breast neoplasms	0.8%	0.0%	1.7%	3.0%	4.0%	1.9%	0.6%	0.0%	1.3%
02.05 Female reproductive neoplasms	1.2%	0.0%	2.5%	0.4%	0.0%	0.7%	1.3%	0.0%	2.6%
02.99 Other and unspecified neoplasms	3.5%	4.3%	2.7%	1.9%	0.4%	3.5%	4.2%	5.1%	3.2%
03.01 Severe anaemia	0.0%	0.0%	0.0%	3.3%	3.5%	3.0%	0.0%	0.0%	0.0%
03.02 Severe malnutrition	0.3%	0.3%	0.3%	0.3%	0.1%	0.5%	0.0%	0.1%	0.0%
03.03 Diabetes mellitus	5.5%	4.1%	7.2%	5.0%	3.9%	6.2%	5.4%	4.6%	6.2%
04.01 Acute cardiac disease	2.8%	3.1%	2.5%	4.2%	4.4%	4.1%	2.2%	1.9%	2.5%
04.02 Stroke	4.8%	3.9%	5.9%	8.8%	7.3%	10.5%	5.7%	3.8%	7.7%
04.99 Other and unspecified cardiac disease	10.6%	8.7%	12.8%	8.5%	6.0%	11.1%	11.9%	8.0%	16.2%
05.01 Chronic obstructive pulmonary disease (COPD)	2.0%	2.4%	1.5%	0.7%	0.9%	0.6%	1.3%	1.7%	0.8%

VA list	Stats SA 2016			Verbal autopsy NCOD Validation 2017/18					
				InterVA-5			Doctor reviewed / IRIS coded		
	Person	Male	Female	Person	Male	Female	Person	Male	Female
05.02 Asthma	0.8%	0.8%	0.8%	0.0%	0.0%	0.1%	1.3%	1.2%	1.3%
06.01 Acute abdomen	0.3%	0.3%	0.3%	0.3%	0.2%	0.4%	0.0%	0.0%	0.0%
06.02 Liver cirrhosis	1.0%	0.8%	0.9%	1.0%	0.9%	1.2%	1.2%	1.1%	1.3%
07.01 Renal failure	1.7%	1.7%	1.7%	1.5%	1.4%	1.6%	0.6%	0.7%	0.5%
08.01 Epilepsy	0.8%	0.9%	0.6%	0.7%	0.9%	0.5%	0.9%	1.2%	0.7%
09.01 Ectopic pregnancy	0.0%	0.0%	0.0%	0.1%	0.0%	0.1%	0.0%	0.0%	0.0%
09.02 Abortion-related death	0.0%	0.0%	0.0%	0.2%	0.0%	0.4%	0.1%	0.0%	0.1%
09.03 Pregnancy-induced hypertension	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
9.04 Obstetric haemorrhage	0.0%	0.0%	0.0%	0.4%	0.0%	0.9%	0.1%	0.0%	0.2%
09.05 Obstructed labour	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
09.06 Pregnancy-related sepsis	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
09.08 Ruptured uterus	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
09.99 Other and unspecified maternal cause	0.3%	0.2%	0.5%	0.0%	0.0%	0.1%	0.1%	0.0%	0.1%
10.01 Prematurity	0.2%	0.2%	0.2%	0.3%	0.2%	0.4%	0.2%	0.1%	0.4%
10.02 Birth asphyxia	0.4%	0.4%	0.3%	0.4%	0.4%	0.4%	0.4%	0.5%	0.4%
10.03 Neonatal pneumonia	0.2%	0.2%	0.2%	0.1%	0.1%	0.2%	0.1%	0.2%	0.1%
10.04 Neonatal sepsis	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.1%	0.1%	0.2%
10.06 Congenital malformation	0.5%	0.5%	0.5%	0.6%	0.6%	0.7%	0.5%	0.4%	0.7%
10.99 Other and unspecified perinatal cause of death	0.8%	0.8%	0.7%	0.1%	0.1%	0.1%	0.3%	0.3%	0.2%
12.01 Road traffic accident	1.4%	2.0%	0.8%	1.6%	2.3%	0.8%	2.6%	3.5%	1.6%
12.02 Other transport accident	0.0%	0.0%	0.0%	0.1%	0.1%	0.0%	0.0%	0.0%	0.0%
12.03 Accidental fall	0.0%	0.1%	0.0%	0.3%	0.4%	0.2%	0.7%	1.0%	0.4%
12.04 Accidental drowning and submersion	0.3%	0.5%	0.1%	0.3%	0.4%	0.1%	0.2%	0.4%	0.0%
12.05 Accidental exposure to smoke, fire and flames	0.5%	0.6%	0.4%	0.3%	0.4%	0.2%	0.2%	0.3%	0.1%
12.06 Contact with venomous animals and plants	0.0%	0.0%	0.0%	0.1%	0.1%	0.0%	0.1%	0.1%	0.0%

VA list	Stats SA 2016			Verbal autopsy NCOD Validation 2017/18					
				InterVA-5			Doctor reviewed / IRIS coded		
	Person	Male	Female	Person	Male	Female	Person	Male	Female
12.07 Accidental poisoning and exposure to noxious substance	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.4%	0.5%	0.2%
12.08 Intentional self-harm	0.1%	0.1%	0.0%	2.0%	2.4%	1.5%	1.2%	1.8%	0.5%
12.09 Assault	1.7%	2.8%	0.4%	7.3%	11.0%	3.2%	4.0%	5.9%	1.9%
12.10 Exposure to force of nature	0.1%	0.1%	0.1%	0.3%	0.3%	0.3%	0.2%	0.4%	0.0%
12.99 Other and unspecified external cause of death	6.9%	10.1%	3.3%	0.2%	0.3%	0.2%	3.3%	4.5%	1.9%
98 Other and unspecified non-communicable disease	10.9%	10.1%	11.8%	0.9%	0.7%	1.0%	5.4%	5.5%	5.4%
99 Cause of death unknown	12.5%	12.0%	13.2%	4.7%	5.0%	4.4%	8.6%	8.8%	8.4%
Total	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**

Causes with large differences between Stats SA proportion and the VA proportion are highlighted.

5. DISCUSSION

5.1 Key findings

- This study has demonstrated that collecting cause of death data using a verbal autopsy nationally was feasible and could provide good quality cause of death information when reviewed by trained doctors. However, recruitment of next of kin through funeral undertakers or Home Affairs offices was extremely challenging.
- The verbal autopsy data collected for this study have demonstrated that HIV/AIDS and TB was measurable by verbal autopsy in a high HIV burden country using either InterVA-5 or physician reviewed. Using physician reviewed verbal autopsy, we found that HIV/AIDS and TB accounted for 28.9% versus 4.9% in the 2016 Stats SA data. In addition, HIV and TB comorbidity is not generally reported by Stats SA (because it required 4-digit ICD coding) but could be identified through the physician reviewed verbal autopsy, however not through InterVA-5.
- A concerning finding revealed by the information from the narratives was the number of HIV positive cases where it was reported that they had been on antiretroviral therapy and had discontinued treatment and then died from HIV-related causes. There was mention of defaulting on treatment in 10.2% of the HIV/AIDS deaths.
- In our experience, the narrative section of the verbal autopsy provided critical information for establishing the cause of death particularly for clinicians and should be at the beginning of the interview. The verbal autopsy questionnaire was very long and needs to be shortened if possible. Specific sections that created confusion for interviewers and respondents included the relationship to the decedent, the injury questions, and the maternal health questions, e.g. 10 women of more than 60 years had a response suggesting that they had been pregnant at the time of or within 42 days of death. More work is needed to fully investigate the performance of the verbal autopsy instruments.
- This study has also demonstrated that it was possible to scan medical and forensic records to provide clear images for review at centralized sites by medical professionals.

5.2 Study limitations

- The main limitation of this study was the poor sample realisation for verbal autopsy of only 15% due to difficulties in getting access to next of kin without contravening the Protection of Personal Information (POPI) Act. In addition, many of the sampled cases fell outside of the target study population due to poor compliance with the inclusion criteria in Phase 1 of data collection. The cause of death patterns presented in this report cannot be assumed to be nationally representative because of the poor sample of verbal autopsy results achieved. However, the sample does have national coverage and are largely consistent with the national burden of disease profile. The potential bias in the verbal autopsy sample is not expected to have a major impact on estimation of the correction factors, the main objective of the project.
- Another limitation was that the pre-test of the method of recruitment was too localized. It was conducted in an area very near to the head office of the fieldwork company contracted for Phase 1 where their staff were familiar to the community. If the pre-test had covered other areas around South Africa, it may have highlighted the challenges of recruitment at an earlier stage.
- This project was a very large national study using methodology that had never been used in South Africa before and made the planning and budgeting difficult.

5.3 Study strengths

- Good quality cause of death data were collected. The fieldwork to conduct the verbal autopsies and collect facility records was very well prepared, conducted, and monitored. The use of experienced interviewers who had previously conducted household surveys resulted in a very low refusal rate for verbal autopsy interviews by the next of kin who had been recruited at the time of death. The interviewers had undergone excellent training in verbal autopsy, and also had the opportunity to conduct practice interviews in suburbs around Tshwane. This has resulted in verbal autopsy narratives and verbal autopsy interview data of good quality.
- Digital data collection tools using KoBoToolbox enabled ongoing monitoring and immediate identification of data quality issues. This quality assurance has ensured good quality data.
- Thorough training of doctors to conduct the reviews of the verbal autopsies and the medical and forensic pathology records and identify the underlying cause of death has resulted in good quality data. Materials from previous trainings for doctors in medical certification provided the basis for the training of study doctors, together with input that was provided by experienced collaborators during pre-testing phases.
- The project has built capacity for cause of death determination which will remain beyond the study.
- The project will enable cause of death validation at a national level.

6. RECOMMENDATIONS

6.1 Link the data with Stats SA data to estimate correction factors

It is critical for the data processing to be completed and as soon as the 2017 Stats SA data becomes available, the data collected in the project will be linked with the CRVS data in Stats SA. The high proportion of HIV/AIDS deaths found in the sample of deaths highlights the importance of estimating correction factors that can assist with providing informative cause of death profiles. Once the linkage has been done, it will be important to assess whether the correction factors need to be weighted to adjust for any potential bias in the data.

6.2 Enhance verbal autopsy tools

Further analysis of the verbal autopsy data that has been collected and processed is needed to consider various issues:

- The correspondence between the doctor reviewed cause of death and Inter-VA5 most probable cause data indicates room for improvement (see Table 16 and Table 17). The symptom-cause relationship in this sample could possibly be used with InterVA and other automated tools to provide an enhanced cause of death profile than the InterVA has done.
- It is desirable to shorten the verbal autopsy questionnaires. The time to complete a questionnaire varied according to the demographic of the decedent. Interviews regarding adult male decedents took the least time. Children and females took the longest, especially where maternal deaths were concerned. When documents such as the death certificate and Road to Health (RTH) cards were available, it increased time spent in the household. Including the time taken for initial contact and consent, a single interview could take 3 hours, depending on the factors mentioned. Regarding the completion time for the questionnaire only by the next of kin/carer, it generally took between 30-45 minutes for a death for an adult male while it took 45 – 60 mins for an adult female or child. The project team could collaborate with the WHO initiative to investigate the performance of all the items in the verbal autopsy questionnaires to assess if there are opportunities for item reduction.

6.3 Train doctors in medical certification

The high quality of the cause of death information provided by the study doctors emphasizes the importance of training doctors in the ICD principles of underlying cause of death and how to complete the medical certificate. The training resources used for this study are currently being adapted into an online training platform that will enable self-learning and assessment linked to Continuing Education Units. Offering Ethics Continuing Education Units would provide an incentive for both public and private doctors to complete the course as these are generally less available. The platform may be evaluated for use in academic settings during medical training (under-graduate and internships), in the public sector during compulsory community service year and when physicians are newly appointed, and in the private sector. This is one opportunity to enhance the quality of cause of death statistics in South Africa.

6.4 Use of verbal autopsy on a national scale

Lessons from this project could be shared with stakeholders of CRVS systems to assess the potential role of using verbal autopsy to complement national statistical information systems. Resource implications could be considered as well as overcoming the challenge of recruitment of next of kin. Although the facility-based data are still to be analysed, it might be important to prioritise collection of such information about the out of facility deaths, particularly child deaths, based on the assumption that the CRVS does not obtain good quality information. It would be valuable to undertake a careful assessment of the optimal use and feasibility of this.

7. REFERENCES

1. World Health Organisation. International statistical classification of diseases and related health problems. Tenth revision. Geneva: WHO; 2011.
2. World Health Organisation. The 2016 WHO verbal autopsy instrument. c2020. https://www.who.int/healthinfo/statistics/verbalautopsystandards/en/ (accessed 12 Feb 2020.
3. Lopez AD, Setel PW. Better health intelligence: a new era for civil registration and vital statistics? *BMC medicine* 2015; **13**: 73.
4. Roth GA, Abate D, Abate KH, et al. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* 2018; **392**(10159): 1736-88.
5. Phillips D, Lozano R, Naghavi M, et al. A composite metric for assessing data on mortality and causes of death: the vital statistics performance index. *Population Health Metrics* 2014; **12**(1): 14.
6. United Nations Statistics Division. Principles and recommendations for a vital statistics system. Statistical Papers, Series M, No. 19, Revision 3. New York: United Nations; 2014.
7. United Nations Statistics Division. Handbook on Civil Registration and Vital Statistics Systems: Management, Operation and Maintenance: Revision 1. New York, USA: United Nations, 2018.
8. Mills SL, Abouzahr C, Lee JKM, M. Rassekh B, Sarpong D. Civil registration and vital statistics (CRVS) for monitoring the Sustainable Development Goals (SDGs). Washington, D.C.: World Bank Group, 2017.
9. United Nations. Principles for a vital statistics system: recommendations for the improvement and standardization of vital statistics. New York: United Nations, 1953.
10. United Nations. Handbook of vital statistics methods. New York: United Nations, 1955.
11. United Nations. Handbook on civil registration and vital statistics systems: Management, operation and maintenance. Studies in Methods. New York: United Nations; 1998.
12. United Nations Population Division. World Mortality Report 2005. New York: United Nations, 2006.
13. United Nations Statistics Division. Revising current international standards on civil registration and vital statistics: proposed outline. New York: United Nations, 2011.
14. United Nations, World Health Organisation. Generating reliable cause-of-death information within a civil registration and vital statistics system: an overview of concepts, standards and implementation. New York: United Nations, 2012.
15. World Health Organization. Annual Epidemiological and Vital Statistics, 1939-1946. Part I: Vital Statistics and Causes of Death; Part II: Causes of Death from Notifiable Diseases. New York, USA: Columbia University Press; 1952.
16. World Health Organisation. World Health Statistics Annual. Geneva: World Health Organization, 1998 & 1989.
17. World Health Organisation. Country health information systems: a review of the current situation and trends. Geneva, Switzerland: World Health Organisation, 2011.
18. World Health Organisation. World Health Statistics 2014. Geneva: World Health Organization., 2014.
19. World Health Organisation. Improving mortality statistics through civil registration and vital statistics systems: guidance for country strategies and partner support. Outcome of a technical meeting, 4 - 5 November 2014. Geneva: World Health Organisation; 2014.
20. The World Bank. World Development Report 1993: Investing in Health. Washington, DC: The World Bank, 1993.
21. World Bank, World Health Organisation. Global Civil Registration and Vital Statistics: Scaling-up Investment Plan 2015-2024. Washington, DC, USA: World Bank Group, 2014.
22. Wang H, Dwyer-Lindgren L, Lofgren KT, et al. Age-specific and sex-specific mortality in 187 countries, 1970-2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet* 2012; **380**(9859): 2071-94.
23. Lozano R, Naghavi M, Foreman K, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet* 2012; **380**(9859): 2095-128.

24. GBD 2016 Causes of Death Collaborators. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet (London, England)* 2017; **390**(10100): 1151-210.
25. United Nations. Transforming Our World: The 2030 Agenda for Sustainable Development. Resolution A/RES/70/1. New York: United Nations General Assembly: United Nations, 2015, Sep 25.
26. United Nations Statsitics Division. SDG Indicators: Global indicator framework for the Sustainable Development Goals and targets of the 2030 Agenda for Sustainable Development. 2019. https://unstats.un.org/sdgs/indicators/indicators-list/ (accessed 08 Oct 2019 2019).
27. Bradshaw D, Nannan N, Groenewald P, et al. Provincial mortality in South Africa, 2000--priority-setting for now and a benchmark for the future. *South African Medical Journal* 2005; **95**(7): 496-503.
28. Msemburi W, Pillay-van Wyk V, Dorrington R, et al. Second national burden of disease study for South Africa: Cause-of-death profile for Eastern Cape, 1997–2012. ISBN: 978-1-928340-08-9. Cape Town, South Africa: South African Medical Research Council, 2016.
29. Groenewald P, Msemburi W, Neethling I, Day C, Tuoane-Nkhazi M, Bradshaw D. Burden of Disease. In: Massyn N, Day C, Barron P, Haynes R, English R, Padarath A, eds. District Health Barometer 2011/12. Durban, South Africa: Health Systems Trust; 2013: 131-50.
30. Bradshaw D, Groenewald P, Bourne DE, et al. Making COD statistics useful for public health at local level in the city of Cape Town. *Bulletin of the World Health Organization* 2006; **84**(3): 211-7C.
31. Groenewald P, Bradshaw D, Daniels J, et al. Local-level mortality surveillance in resource-limited settings: a case study of Cape Town highlights disparities in health. *Bulletin of the World Health Organization* 2010; **88**(6): 444-51.
32. Joubert JD. Availability and quailty of vital registration mortality data in South Africa: A national and surveillance site perspective. PhD thesis. Brisbane, Queensland, Australia: The University of Queensland; 2015.
33. Joubert J, Rao C, Bradshaw D, Vos T, Lopez AD. Evaluating the quality of national mortality statistics from civil registration in South Africa, 1997–2007. *PLoS ONE* 2013; **8**(5): e64592.
34. Republic of South Africa. Births and Deaths Registration (Act 51 of 1992). Government Gazette No. 13953, Vol. 323. Cape Town, South Africa: Government Printer; 1992.
35. Joubert J, Rao C, Bradshaw D, Dorrington RE, Vos T, Lopez AD. Characteristics, availability and uses of vital registration and other mortality data sources in post-democracy South Africa. *Global Health Action* 2012; **5**(19263).
36. Dorrington RE, Bourne D, Bradshaw D, Laubscher R, Timaeus IM. The impact of HIV/AIDS on adult mortality in South Africa. Tygerberg, Cape Town: South African Medical Research Council, 2001.
37. Bradshaw D, Schneider M, Dorrington RE, Bourne DE, Laubscher R. South African cause-of-death profile in transition - 1996 and future trends. *S Afr Med J* 2002; **92**(8 I): 618-23.
38. Bradshaw D, Laubscher R, Dorrington RE, Bourne DE, Timæus IM. Unabated rise in number of adult deaths in South Africa. *South African Medical Journal* 2004; **94**(4): 278-9.
39. Botha JL, Bradshaw D. African vital statistics - A black hole? *South African Medical Journal* 1985; **67**(24): 977-81.
40. Bradshaw D, Dorrington RE, Sitas F. The level of mortality in South Africa in 1985 - What does it tell us about health? *South African Medical Journal* 1992; **82**(4): 237-40.
41. Shisana S, Bradshaw D. Births and Deaths Registration Act, 1992 [23]. *South African Medical Journal* 1993; **83**(2): 147.
42. Bradshaw D, Kielkowski D, Sitas F. New birth and death registration forms - A foundation for the future, a challenge for health workers? *South African Medical Journal* 1998; **88**(8): 971-4.
43. Dorrington RE, Bradshaw D, Wegner T. Estimates of the Level and Shape of Mortality Rates in South Africa around 1985 and 1990 derived by applying indirect demographic techniques to reported deaths. MRC Technical Report. Cape Town: South African Medical Research Council; 1999.

44. Timaeus I, Dorrington RE, Bradshaw B, Nannan N, Bourne D. Adult mortality in South Africa, 1980-2000: From apartheid to AIDS. Population Association of America Annual Meeting. Washington, DC; 2001. p. 29-31.
45. Bradshaw D, Groenewald P, Laubscher R, et al. Initial burden of disease estimates for South Africa, 2000. *South African Medical Journal* 2003; **93**(9): 682-8.
46. Bchir A, Bhutta Z, Binka F, et al. Better health statistics are possible. *The Lancet* 2006; **367**(9506): 190-3.
47. Nojilana B, Groenewald P, Bradshaw D, Reagon G. Quality of cause of death certification at an academic hospital in Cape Town, South Africa. *South African Medical Journal* 2009; **99**(9): 648-52.
48. Dorrington RE, Moultrie TA, Timæus IM. Estimation of mortality using the South African Census 2001 data. Cape Town: Centre for Actuarial Research, University of Cape Town, 2004.
49. Bah S. Multiple forces working in unison: the case of rapid improvement of vital statistics in South Africa post-1996. *World Health & Population* 2009; **11**(1): 50-9.
50. Pillay-Van Wyk V, Laubscher R, Msemburi W, et al. Second South African National Burden of Disease Study: Data cleaning, validation and SA NBD list. Cape Town, South Africa: South African Medical Research Council, 2014.
51. Garenne M, Collinson MA, Kabudula CW, Gómez-Olivé FX, Kahn K, Tollman S. Improving completeness of birth and death registration in rural Africa. *The Lancet Global Health* 2016; **4**(9): e604-e5.
52. Johansson L, Pavillon G, Pelikan L, Weber S. Iris automated coding system for causes of death: User's reference manual. Version 4.1.3. German Institute of Medical Documentation and Information (DIMDI): IRIS Institute; 2012.
53. World Health Organization. International Statistical Classification of Diseases and Related Health Problems. Tenth Revision. Instruction Manual. Geneva, Switzerland: World Health Organization; 2004.
54. World Health Organization. International statistical classification of diseases and related health problems. 10th Revision. Instruction manual. Geneva: World Health Organization; 2011.
55. Statistics South Africa. Mortality and causes of death in South Africa, 2016: findings from death notification. Statistical release P0309.3. Pretoria, South Africa: Statistics South Africa, 2018.
56. Pillay-van Wyk V, Bradshaw D, Groenewald P, Laubscher R. Improving the quality of medical certification of cause of death: the time is now! *S Afr Med J* 2011; **101**(9): 626.
57. Massyn N, Day C, Peer N, Padarath A, Barron P, English R, editors. District Health Barometer 2013/14. Durban: Health Systems Trust; 2014.
58. Westwood AT. Childhood deaths due to HIV--the role of the new death certificate. *South African Medical Journal* 2000; **90**(9): 877.
59. Grandin W, Westwood T, Lagerdien K, King MS. Deaths at Red Cross Children's Hospital, Cape Town 1999-2003--a study of death notification forms. *South African Medical Journal* 2006; **96**(9 Pt 2): 964-8.
60. Yudkin PL, Burger EH, Bradshaw D, Groenewald P, Ward AM, Volmink J. Deaths caused by HIV disease under-reported in South Africa. *AIDS* 2009; **23**(12): 1600-2.
61. Burger EH, Groenewald P, Bradshaw D, Ward AM, Yudkin PL, Volmink J. Validation study of cause of death statistics in Cape Town, South Africa, found poor agreement. *Journal of clinical epidemiology* 2012; **65**(3): 309-16.
62. Joubert J, Bradshaw D, Kabudula C, et al. Record-linkage comparison of verbal autopsy and routine civil registration death certification in rural north-east South Africa: 2006-09. *International journal of epidemiology* 2014; **43**(6): 1945-58.
63. Groenewald P, Bradshaw D, Dorrington R, Bourne D, Laubscher R, Nannan N. Identifying deaths from AIDS in South Africa: An update [4]. *AIDS* 2005; **19**(7): 744-5.
64. Birnbaum JK, Murray CJL, Lozano R. Exposing misclassified HIV/AIDS deaths in South Africa. *Bull World Health Organ* 2011; **89**(4): 278-85.
65. Matzopoulos R, Prinsloo M, Pillay-van Wyk V, et al. Injury-related mortality in South Africa: a retrospective descriptive study of postmortem investigations. *Bull World Health Organ* 2015; **93**(5): 303-13.
66. Johansson LA, Westerling R, Rosenberg HM. Methodology of studies evaluating death certificate accuracy were flawed. *Journal of clinical epidemiology* 2006; **59**(2): 125-31.
67. Lozano R, Lopez AD, Atkinson C, et al. Performance of physician-certified verbal autopsies: multisite validation study using clinical diagnostic gold standards. *Population Health Metrics* 2011; **9**: 32.

68. von Stillfried S, Isfort P, Knuchel-Clarke R. [Postmortem imaging procedures: Experiences and perspectives] Postmortale bildgebende Verfahren: Erfahrungen und Ausblicke. *Der Pathologe* 2017; **38**(5): 412-5.
69. Castillo P, Ussene E, Ismail MR, et al. Pathological methods applied to the investigation of causes of death in developing countries: Minimally invasive autopsy approach. *PLoS One* 2015; **10**(6): e0132057.
70. Murray CJ, Lopez AD, Black R, Ahuja R, Ali SM, Baqui A. Population Health Metrics Research Consortium gold standard verbal autopsy validation study: design, implementation, and development of analysis datasets. *Popul Health Metr* 2011; **9**.
71. Rao C, Yang G, Hu J, Ma J, Xia W, Lopez AD. Validation of cause-of-death statistics in urban China. *International journal of epidemiology* 2007; **36**(3): 642-51.
72. Leth P. Computerized Tomography Used as a Routine Procedure at Postmortem Investigations. *The American journal of forensic medicine and pathology* 2009; **30**: 219-22.
73. Tal S, Berkovitz N, Gottlieb P, Zaitsev K. Acceptance of forensic imaging in Israel. *The Israel Medical Association Journal : IMAJ* 2015; **17**(3): 141-4.
74. Lewis C, Latif Z, Hill M, et al. "We might get a lot more families who will agree": Muslim and Jewish perspectives on less invasive perinatal and paediatric autopsy. *PLoS One* 2018; **13**(8): e0202023.
75. Milner DA, Jr., Dzamalala CP, Liomba NG, Molyneux ME, Taylor TE. Sampling of supraorbital brain tissue after death: improving on the clinical diagnosis of cerebral malaria. *The Journal of infectious diseases* 2005; **191**(5): 805-8.
76. Weustink AC, Hunink MG, van Dijke CF, Renken NS, Krestin GP, Oosterhuis JW. Minimally invasive autopsy: an alternative to conventional autopsy? *Radiology* 2009; **250**(3): 897-904.
77. Martínez MJ, Massora S, Mandomando I, et al. Infectious cause of death determination using minimally invasive autopsies in developing countries. *Diagnostic Microbiology and Infectious Disease* 2016; **84**(1): 80-6.
78. Johansson LA. Targeting non-obvious errors in death certificates. Uppsala, Sweden: Uppsala University; 2008.
79. Rampatige R, Mikkelsen L, Prado BH, Riley I, Lopez AD. Assessing the reliability of hospital-based cause-of-death statistics: evidence-based guidelines for country application. Brisbane, Australia: Health Information Systems Knowledge Hub, School of Population Health, The University of Queensland, 2013.
80. Statistics South Africa. Mortality and causes of death in South Africa, 1997 - 2013. Dataset produced and provided by Statistics South Africa, Pretoria, South Africa, based on information extracted from official civil registration Death Notification Forms submitted for death registration to the Department of Home Affairs, for deaths occurring in the years 1997 - 2013;; 2014.
81. Kahn K, Tollman SM, Garenne M, Gear JSS. Validation and application of verbal autopsies in a rural area of South Africa. *Tropical Medicine & International Health* 2000; **5**(11): 824-31.
82. Bradshaw D, Msemburi W, Dorrington R, et al. HIV/AIDS in South Africa: how many people died from the disease between 1997 and 2010? *Aids* 2015.
83. Naghavi M, Wang H, Lozano R, et al. Global, regional, and national age–sex specific all-cause and cause-specific mortality for 240 causes of death, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet* 2014; **385**(9963): 117-71.
84. Statistics South Africa. Census 2011. Statistics Release P0304.1. Pretoria: Statistics South Africa, 2012.
85. Rao C, Porapakkham Y, Pattaraarchachai J, Polprasert W, Swampunyalert N, Lopez AD. Verifying causes of death in Thailand: rationale and methods for empirical investigation. *Population Health Metrics* 2010; **8**(1): 11.
86. Pillay-van Wyk V, Msemburi W, Laubscher R, et al. Mortality trends and differentials in South Africa from 1997 to 2012: second National Burden of Disease Study. *The Lancet Global Health* 2016; **4**(9): e642-e53.
87. Lwanga SK, Lemishow S. Sample size determination in health studies. A practical manual. Geneva: World Health Organisation; 1991.
88. Harvard Humanitarian Initiative. KoBoToolbox: Data Collection Tools for Challenging Environments. Available at: http://www.kobotoolbox.org/; .
89. Open Data Kit. https://opendatakit.org/ Powered by Jekyll and Minimal Mistakes, Accesssed 11 Dec 2019.
90. Clear Scanner. Available: https://clear-scanner-free-pdf-scans.en.softonic.com/android Accessed: 11 Dec 2019.

91. Microsoft Teams. (Version release Android 1416/1.0.0.2019101801) [Desktop application software] Available: http://products.office.com/en-us/microsoft-teams/group-chat-software Accessed: 05 Dec 2019.
92. Byass P, Hussain-Alkhateeb L, D'Ambruoso L, et al. An integrated approach to processing WHO-2016 verbal autopsy data: the InterVA-5 model. *BMC medicine* 2019; **17**(1): 102.
93. Hussain-Alkhateeb L, D'Ambruoso L, Tollman S, et al. Enhancing the value of mortality data for health systems: adding Circumstances Of Mortality CATegories (COMCATs) to deaths investigated by verbal autopsy. *Glob Health Action* 2019; **12**(1): 1680068.
94. Iris Institute. Iris – Automated coding system for causes of death. Heidelberg, Germany: German Institute of Medical Documentation and Information (DIMDI), 2018.
95. Mobenzi. Available: https://www.mobenzi.com/ Accessed: 09 Dec 2019.
96. Dropbox Business. Available: https://www.dropbox.com/business Accessed: 11 Dec 2019. .
97. Chapman WB. Validate South African ID Number with Excel Formula. 2019. https://wa.rwick.com/author/admin/ (accessed Accessed: 09 Dec 2019.
98. Dorrington RE, Bradshaw D, Laubscher R, Nannan N. Rapid mortality surveillance report 2018. Cape Town: South African Medical Research Council, 2020.
99. Mikkelsen L, Lopez AD. Guidance for assessing and interpreting the quality of mortatlity data using ANACONDA. Melbourne, Australia: Bloomberg Philanthropies Data for Health Initiative, and the Civil Registration and Vital Statistics Improvement Group of the Melbourne University, 2017.
100. Mikkelsen L, Richards N, Lopez AD. Redefining 'garbage codes' for public health policy: Report on the expert group meeting, 27 – 28 February 2017. CRVS technical outcome series. Melbourne, Australia: Bloomberg Philanthropies Data for Health Initiative, and the Civil Registration and Vital Statistics Improvement project of the University of Melbourne, 2018.
101. Republic of South Africa. Protection of Personal Information (POPI) (Act 4 of 2013). Government Gazette No. 37544, Vol. 581. Cape Town, South Africa: Government Printer; 2013.
102. Republic of South Africa. Statistics Act (Act No. 6 of 1999). Government Gazette No. 19957, Vol. 406, 21 April 1999. Cape Town: Government Printer; 1999.

8. ANNEXURE

8.1 Objectives of SA NCOD Validation Project 2017/18

The study has three interrelated objectives which each have their own more detailed sub-objectives:

1. To verify causes of death reported on death notification forms in a nationally-representative sample of deaths occurring within and outside health facilities.
    a. For deaths occurring in health facilities, agreement between the underlying cause of death reported on the DHA 1663 and the underlying cause of death based on medical records will be measured.
    b. For deaths occurring outside health facilities, the agreement between the underlying cause of death reported on the DHA 1663 and the underlying cause of death obtained from an interviewer-administered household VA will be measured.
    c. For deaths requiring a forensic investigation, the agreement between the underlying cause of death (external or natural) reported on the DHA 1663 and the underlying cause of death (external or natural) reported in forensic records will be measured.
    d. To check whether decedents were recorded in appropriate death registers (e.g., cancer register, Tier.net (ie HIV register) or the TB register).

2. To derive correction factors to adjust cause-specific mortality data from vital registration according to reference diagnoses at national, provincial, and district level.
    a. Correction factors for reference diagnoses will be derived from national sample data.
    b. The nationally derived correction factors for reference diagnoses will be applied to cause of death profiles from vital registration data at national, provincial, and district level.

3. To design and test a standardized methodology for household VA for deaths occurring outside health facilities, with a view towards broader implementation within the routine CVRS system.
    a. The agreement between physician coded VA underlying cause of death and the underlying cause of death obtained from medical and forensic records, will be measured for deaths occurring in health facilities and those requiring a forensic investigation.
    b. The agreement between the cause-specific mortality fraction (CSMF) produced through automated coding of VA and CSMF from medical and forensic records, will be measured for deaths occurring in health facilities and those requiring a forensic investigation.
    c. The feasibility and community acceptability of implementing VA as a routine part of the CVRS system will be assessed based upon interviewer experience in the field.

8.2 Geographic sampling frame (health sub-districts)

Rank of sub-districts by poverty headcount in each Tertile	Province	Sub-district	Municipal sub-category	2011 poverty headcount	2011 Population	Estimated 2013 deaths
1	**Eastern Cape**	261 EC101: Camdeboo	B3	2,8%	50 993	443
2	**Eastern Cape**	278 EC131: Inxuba Yethemba	B3	3,0%	65 560	722
3	**Eastern Cape**	269 EC109: Kou-Kamma	B3	3,2%	40 663	354
4	**Eastern Cape**	Nelson Mandela Bay C	A	3,45%	458364	2784
5	**Eastern Cape**	263+267: Ikwezi+Baviaans	B3	3,5%	28298	246
6	**Eastern Cape**	262 EC102: Blue Crane Route	B3	4,7%	36 002	313
7	**Eastern Cape**	264 EC104: Makana	B2	5,1%	80 390	699
8	**Eastern Cape**	289 EC144: Gariep	B3	5,2%	33 677	432
9	**Eastern Cape**	Nelson Mandela Bay A	A	5,33%	393021	2387
10	**Eastern Cape**	280 EC133: Inkwanca	B3	5,5%	21 971	242
11	**Eastern Cape**	266 EC106: Sundays River Valley	B3	5,7%	54 504	474
12	**Eastern Cape**	Nelson Mandela Bay B	A	5,79%	300738	1826
13	**Eastern Cape**	268 EC108: Kouga	B3	5,9%	98 558	857
1	**Eastern Cape**	277 EC128: Nxuba	B3	6,4%	24 264	336
2	**Eastern Cape**	281 EC134: Lukanji	B2	7,1%	190 723	2100
3	**Eastern Cape**	265 EC105: Ndlambe	B3	7,4%	61 176	532
4	**Eastern Cape**	288 EC143: Maletswai	B3	8,9%	43 800	562
5	**Eastern Cape**	260 BUF: Buffalo City	A	9,3%	755 200	8043
6	**Eastern Cape**	279 EC132: Tsolwana	B3	11,7%	33 281	367
7	**Eastern Cape**	272 EC123: Great Kei	B3	12,2%	38 991	539
8	**Eastern Cape**	276 EC127: Nkonkobe	B3	13,9%	127 115	1759
9	**Eastern Cape**	285 EC138: Sakhisizwe	B3	14,3%	63 582	700
10	**Eastern Cape**	273 EC124: Amahlathi	B3	14,3%	122 778	1699
11	**Eastern Cape**	287 EC142: Senqu	B4	14,5%	134 150	1721
12	**Eastern Cape**	274 EC126: Ngqushwa	B4	14,6%	72 190	999
13	**Eastern Cape**	294 EC157: King Sabata Dalindyebo	B2	15,5%	451 710	4279
1	**Eastern Cape**	283 EC136: Emalahleni	B4	17,2%	119 460	1316
2	**Eastern Cape**	271 EC122: Mnquma	B4	20,7%	252 390	3492
3	**Eastern Cape**	292 EC155: Nyandeni	B4	21,2%	290 390	2751
4	**Eastern Cape**	293 EC156: Mhlontlo	B4	21,4%	188 226	1783
5	**Eastern Cape**	295 EC441: Matatiele	B3	22,4%	203 843	1501
6	**Eastern Cape**	282 EC135: Intsika Yethu	B4	22,9%	145 372	1601
7	**Eastern Cape**	286 EC141: Elundini	B4	24,7%	138 141	1773
8	**Eastern Cape**	297 EC443: Mbizana	B4	25,0%	281 905	2076

Rank of sub-districts by poverty headcount in each Tertile	Province	Sub-district	Municipal sub-category	2011 poverty headcount	2011 Population	Estimated 2013 deaths
9	**Eastern Cape**	296 EC442: Umzimvubu	B4	25,3%	191 620	1411
10	**Eastern Cape**	270 EC121: Mbhashe	B4	25,6%	254 909	3526
11	**Eastern Cape**	290 EC153: Ngquza Hill	B4	27,3%	278 481	2638
12	**Eastern Cape**	284 EC137: Engcobo	B4	27,4%	155 513	1713
13	**Eastern Cape**	291 EC154: Port St Johns	B4	28,2%	156 136	1479
14	**Eastern Cape**	298 EC444: Ntabankulu	B4	33,6%	123 976	913
1	**Free State**	475 FS201: Moqhaka	B2	2,7%	160 532	1676
2	**Free State**	461 FS162: Kopanong	B3	3,8%	49 171	885
3	**Free State**	463 FS164: Naledi	B3	4,4%	24 314	438
4	**Free State**	474 FS196: Mantsopa	B3	4,6%	51 056	639
5	**Free State**	477 FS203: Ngwathe	B3	4,7%	120 520	1258
6	**Free State**	466 FS183: Tswelopele	B3	4,8%	47 625	547
7	**Free State**	499 MAN: Mangaung	A	4,8%	747 431	8968
1	**Free State**	478 FS204: Metsimaholo	B2	5,1%	149 108	1557
2	**Free State**	464 FS181: Masilonyana	B3	5,3%	63 334	728
3	**Free State**	467 FS184: Matjhabeng	B1	5,5%	406 461	4669
4	**Free State**	460 FS161: Letsemeng	B3	5,6%	38 628	695
5	**Free State**	468 FS185: Nala	B3	5,6%	81 220	933
7	**Free State**	462 FS163: Mohokare	B3	6,2%	34 146	614
1	**Free State**	471 FS193: Nketoana	B3	6,3%	60 324	755
2	**Free State**	469 FS191: Setsoto	B3	6,6%	112 597	1410
3	**Free State**	479 FS205: Mafube	B3	6,8%	57 876	604
4	**Free State**	465 FS182: Tokologo	B3	7,7%	28 986	333
5	**Free State**	472 FS194: Maluti a Phofung	B3	7,9%	335 784	4205
6	**Free State**	473 FS195: Phumelela	B3	8,5%	47 772	598
1	**Gauteng**	Johannesburg B Health Sub-District		0,85%	332613	2210
2	**Gauteng**	Johannesburg F Health Sub-District		1,03%	608556	4044
3	**Gauteng**	Johannesburg E Health Sub-District		1,89%	469470	3119
4	**Gauteng**	Johannesburg D Health Sub-District		2,21%	1098963	7302
5	**Gauteng**	Tshwane 3 Health Sub-District		2,44%	585153	4124
6	**Gauteng**	760 GT421: Emfuleni	B1	3,4%	721 663	8239
7	**Gauteng**	Tshwane 4 health Sub-District		3,91%	379347	2674
8	**Gauteng**	Johannesburg C Health Sub-District		4,29%	627831	4172
9	**Gauteng**	Tshwane 6 Health Sub-District		4,52%	605565	4268

Rank of sub-districts by poverty headcount in each Tertile	Province	Sub-district	Municipal sub-category	2011 poverty headcount	2011 Population	Estimated 2013 deaths
1	**Gauteng**	Tshwane 1 Health Sub-District		4,52%	811575	5720
2	**Gauteng**	762 GT423: Lesedi	B3	4,8%	99 520	1136
3	**Gauteng**	Tshwane 2 Health Sub-District		4,82%	339180	2390
4	**Gauteng**	764 GT482: Randfontein	B2	4,9%	149 286	1768
5	**Gauteng**	Ekurhuleni South 1		5,28%	537078	4401
6	**Gauteng**	Ekurhuleni North 2		5,51%	541419	4437
7	**Gauteng**	Ekurhuleni North 1		5,66%	638508	5232
8	**Gauteng**	763 GT481: Mogale City	B1	5,8%	362 422	4291
9	**Gauteng**	Ekurhuleni South 2		6,26%	568680	4660
1	**Gauteng**	Johannesburg A Health Sub-District		6,41%	638112	4240
2	**Gauteng**	761 GT422: Midvaal	B2	6,5%	95 301	1088
3	**Gauteng**	Tshwane 7 Health Sub-District		6,56%	109767	774
4	**Gauteng**	Johannesburg G Health Sub-District		6,61%	659295	4381
5	**Gauteng**	Ekurhuleni East 1		7,06%	559422	4584
6	**Gauteng**	Ekurhuleni East 2		7,14%	333354	2732
7	**Gauteng**	Tshwane 5 Health Sub-District		7,53%	99813	703
8	**Gauteng**	766 GT484: Merafong City	B2	8,5%	197 520	2339
9	**Gauteng**	765 GT483: Westonaria	B2	15,4%	111 767	1323
1	**KwaZulu-Natal**	538 KZN282: uMhlathuze	B1	4,1%	334 459	2937
2	**KwaZulu-Natal**	524 KZN252: Newcastle	B1	5,5%	363 236	3642
3	**KwaZulu-Natal**	563 KZN222: uMngeni	B2	5,7%	92 710	956
4	**KwaZulu-Natal**	566 KZN225: The Msunduzi	B1	5,9%	618 536	6379
5	**KwaZulu-Natal**	eThekwini West		5,99%	783978	3868
6	**KwaZulu-Natal**	eThekwini South		6,54%	1543683	7616
7	**KwaZulu-Natal**	eThekwini North		6,66%	1114692	5500
8	**KwaZulu-Natal**	574 KZN241: Endumeni	B3	7,3%	64 862	588
9	**KwaZulu-Natal**	506 KZN216: Hibiscus Coast	B2	8,0%	256 135	3153
10	**KwaZulu-Natal**	514 KZN232: Emnambithi/Ladysmith	B2	8,1%	237 437	2372
11	**KwaZulu-Natal**	592 KZN292: KwaDukuza	B2	8,6%	231 187	1882
12	**KwaZulu-Natal**	591 KZN291: Mandeni	B4	8,8%	138 078	1124
13	**KwaZulu-Natal**	596 KZN433: Greater Kokstad	B2	9,2%	65 981	822
14	**KwaZulu-Natal**	587 KZN281: Mfolozi	B4	10,0%	122 889	1079
15	**KwaZulu-Natal**	562 KZN221: uMshwathi	B4	10,6%	106 374	1097
16	**KwaZulu-Natal**	564 KZN223: Mpofana	B3	10,8%	38 103	393
17	**KwaZulu-Natal**	595 KZN432: Kwa Sani	B4	10,9%	12 898	161

Rank of sub-districts by poverty headcount in each Tertile	Province	Sub-district	Municipal sub-category	2011 poverty headcount	2011 Population	Estimated 2013 deaths
1	**KwaZulu-Natal**	529 KZN263: Abaqulusi	B3	11,2%	211 060	1719
2	**KwaZulu-Natal**	526 KZN254: Dannhauser	B4	11,6%	102 161	1024
3	**KwaZulu-Natal**	586 KZN275: Mtubatuba	B3	11,7%	175 425	1221
4	**KwaZulu-Natal**	581 KZN266: Ulundi	B4	12,4%	188 317	1534
5	**KwaZulu-Natal**	579 KZN262: UPhongolo	B4	12,5%	127 238	1036
6	**KwaZulu-Natal**	568 KZN227: Richmond	B4	13,0%	65 793	679
7	**KwaZulu-Natal**	578 KZN261: eDumbe	B3	13,4%	82 053	668
8	**KwaZulu-Natal**	561 KZN212: Umdoni	B2	13,8%	78 875	971
9	**KwaZulu-Natal**	565 KZN224: Impendle	B4	14,2%	33 105	341
10	**KwaZulu-Natal**	567 KZN226: Mkhambathini	B3	14,8%	63 142	651
11	**KwaZulu-Natal**	505 KZN215: Ezingoleni	B4	15,0%	52 540	647
12	**KwaZulu-Natal**	580 KZN265: Nongoma	B4	15,3%	194 908	1587
13	**KwaZulu-Natal**	570 KZN234: Umtshezi	B3	15,5%	83 153	831
14	**KwaZulu-Natal**	589 KZN284: uMlalazi	B4	15,6%	213 601	1876
15	**KwaZulu-Natal**	573 KZN236: Imbabazane	B4	16,1%	113 073	1129
16	**KwaZulu-Natal**	585 KZN274: Hlabisa	B4	16,2%	71 925	501
17	**KwaZulu-Natal**	590 KZN285: Mthonjaneni	B3	16,4%	47 818	420
18	**KwaZulu-Natal**	588 KZN283: Ntambanana	B4	16,9%	74 336	653
1	**KwaZulu-Natal**	504KZN214: UMuziwabantu	B3	17,4%	96 556	1189
2	**KwaZulu-Natal**	584KZN273: The Big 5 False Bay	B3	17,6%	35 258	245
3	**KwaZulu-Natal**	569 KZN233: Indaka	B4	18,3%	103 116	1030
4	**KwaZulu-Natal**	571 KZN235: Okhahlamba	B4	18,7%	132 068	1319
5	**KwaZulu-Natal**	575 KZN242: Nqutu	B4	19,5%	165 307	1499
6	**KwaZulu-Natal**	577 KZN245: Umvoti	B3	19,9%	103 093	935
7	**KwaZulu-Natal**	525 KZN253: Emadlangeni	B3	21,4%	34 442	345
8	**KwaZulu-Natal**	594 KZN431: Ingwe	B4	21,4%	100 548	1253
9	**KwaZulu-Natal**	593 KZN293: Ndwedwe	B4	21,7%	140 820	1146
10	**KwaZulu-Natal**	597 KZN434: Ubuhlebezwe	B4	21,7%	101 691	1267
11	**KwaZulu-Natal**	598 KZN435: Umzimkhulu	B4	22,2%	180 302	2247
12	**KwaZulu-Natal**	583 KZN272: Jozini	B4	22,3%	186 502	1298
13	**KwaZulu-Natal**	503 KZN213: Umzumbe	B4	22,8%	160 975	1981
14	**KwaZulu-Natal**	542 KZN286: Nkandla	B4	24,2%	114 416	1005
15	**KwaZulu-Natal**	546 KZN294: Maphumulo	B4	25,4%	96 724	787
16	**KwaZulu-Natal**	560 KZN211: Vulamehlo	B4	29,0%	77 403	953
17	**KwaZulu-Natal**	582 KZN271: Umhlabuyalingana	B4	29,5%	156 736	1091
18	**KwaZulu-Natal**	576 KZN244: Msinga	B4	37,2%	177 577	1610

Rank of sub-districts by poverty headcount in each Tertile	Province	Sub-district	Municipal sub-category	2011 poverty headcount	2011 Population	Estimated 2013 deaths
1	**Limpopo**	976 LIM355: Lepele-Nkumpi	B4	14,6%	230 350	2305
2	**Limpopo**	979 LIM364: Mookgopong	B3	17,2%	35 640	275
3	**Limpopo**	960 LIM331: Greater Giyani	B4	17,4%	244 217	2018
4	**Limpopo**	964 LIM335: Maruleng	B4	18,2%	94 857	784
5	**Limpopo**	974 LIM354: Polokwane	B1	20,1%	628 999	6294
6	**Limpopo**	969 LIM351: Blouberg	B4	20,7%	162 629	1627
7	**Limpopo**	970 LIM352: Aganang	B4	21,2%	131 164	1312
8	**Limpopo**	973 LIM353: Molemole	B4	21,4%	108 321	1084
1	**Limpopo**	962 LIM333: Greater Tzaneen	B4	21,4%	390 095	3223
2	**Limpopo**	961 LIM332: Greater Letaba	B4	22,2%	212 701	1757
3	**Limpopo**	965 LIM342: Mutale	B4	22,4%	91 870	668
4	**Limpopo**	966 LIM343: Thulamela	B4	22,8%	618 462	4500
5	**Limpopo**	987 LIM475: Greater Tubatse	B4	22,9%	335 676	3009
6	**Limpopo**	963 LIM334: Ba-Phalaborwa	B3	24,2%	150 637	1245
7	**Limpopo**	967 LIM341: Musina	B3	24,7%	68 359	497
8	**Limpopo**	981 LIM366: Bela-Bela	B3	25,0%	66 500	513
1	**Limpopo**	985 LIM473: Makhuduthamaga	B4	25,3%	274 358	2460
2	**Limpopo**	968 LIM344: Makhado	B4	25,4%	516 031	3754
3	**Limpopo**	977 LIM361: Thabazimbi	B3	25,6%	85 234	658
4	**Limpopo**	978 LIM362: Lephalale	B3	27,3%	115 767	893
5	**Limpopo**	983 LIM471: Ephraim Mogale	B4	27,4%	123 648	1109
6	**Limpopo**	984 LIM472: Elias Motsoaledi	B4	28,2%	249 363	2236
7	**Limpopo**	980 LIM365: Modimolle	B3	29,5%	68 513	529
8	**Limpopo**	982 LIM367: Mogalakwena	B2	33,6%	307 682	2374
9	**Limpopo**	986 LIM474: Fetakgomo	B4	37,2%	93 795	841
1	**Mpumalanga**	869 MP313: Steve Tshwete	B1	4,3%	229 831	1798
2	**Mpumalanga**	864 MP305: Lekwa	B3	4,5%	115 662	1047
3	**Mpumalanga**	866 MP307: Govan Mbeki	B1	4,5%	294 538	2666
4	**Mpumalanga**	871 MP315: Thembisile	B4	5,6%	310 458	2428
5	**Mpumalanga**	872 MP316: Dr JS Moroka	B4	6,0%	249 705	1953
6	**Mpumalanga**	874 MP322: Mbombela	B1	6,0%	588 794	4995

Rank of sub-districts by poverty headcount in each Tertile	Province	Sub-district	Municipal sub-category	2011 poverty headcount	2011 Population	Estimated 2013 deaths
1	**Mpumalanga**	870 MP314: Emakhazeni	B2	6,4%	47 216	369
2	**Mpumalanga**	873 MP321: Thaba Chweu	B3	6,6%	98 387	835
3	**Mpumalanga**	867 MP311: Victor Khanye	B3	6,9%	75 452	590
4	**Mpumalanga**	868 MP312: Emalahleni	B1	8,0%	395 466	3093
5	**Mpumalanga**	865 MP306: Dipaleseng	B3	8,3%	42 390	384
6	**Mpumalanga**	875 MP323: Umjindi	B3	9,1%	67 156	570
1	**Mpumalanga**	863 MP304: Pixley Ka Seme	B3	9,2%	83 235	753
2	**Mpumalanga**	861 MP302: Msukaligwa	B2	9,2%	149 377	1352
3	**Mpumalanga**	876 MP324: Nkomazi	B4	10,4%	393 030	3334
4	**Mpumalanga**	860 MP301: Albert Luthuli	B4	10,9%	186 010	1684
5	**Mpumalanga**	877 MP325: Bushbuckridge	B4	11,8%	541 248	4592
6	**Mpumalanga**	862 MP303: Mkhondo	B3	15,8%	171 982	1557
1	**North West**	676 NW402: Tlokwe City Council	B1	3,9%	162 762	1814
2	**North West**	677 NW403: City of Matlosana	B1	4,6%	398 676	4443
3	**North West**	673 NW396: Lekwa-Teemane	B3	5,1%	53 248	631
4	**North West**	662 NW373: Rustenburg	B1	7,2%	549 575	4605
5	**North West**	660 NW371: Moretele	B4	7,9%	186 947	1567
6	**North West**	678 NW404: Maquassi Hills	B3	8,1%	77 794	867
1	**North West**	663 NW374: Kgetlengrivier	B3	8,2%	51 049	428
2	**North West**	664 NW375: Moses Kotane	B4	8,3%	242 554	2033
3	**North West**	661 NW372: Madibeng	B1	9,5%	477 381	4000
4	**North West**	671 NW393: Mamusa	B3	10,1%	60 355	715
5	**North West**	667 NW383: Mafikeng	B2	10,6%	291 527	3159
6	**North West**	670 NW392: Naledi	B3	10,7%	66 781	791
1	**North West**	675 NW401: Ventersdorp	B3	11,4%	56 702	632
2	**North West**	668 NW384: Ditsobotla	B3	11,6%	168 902	1830
3	**North West**	669 NW385: Ramotshere Moiloa	B3	13,2%	150 713	1633
4	**North West**	666 NW382: Tswaing	B3	13,4%	124 218	1346
5	**North West**	665 NW381: Ratlou	B4	16,6%	107 339	1163
6	**North West**	672 NW394: Greater Taung	B4	16,6%	177 642	2104
7	**North West**	674 NW397: Kagisano/Molopo	B4	17,0%	105 789	1253

Rank of sub-districts by poverty headcount in each Tertile	Province	Sub-district	Municipal sub-category	2011 poverty headcount	2011 Population	Estimated 2013 deaths
1	**Northern Cape**	366 NC065: Hantam	B3	2,3%	21 578	193
2	**Northern Cape**	364 NC062: Nama Khoi	B3	2,5%	47 041	420
3	**Northern Cape**	362 NC453: Gamagara	B3	2,5%	41 617	467
4	**Northern Cape**	382 NC086: Kgatelopele	B3	2,6%	18 687	219
5	**Northern Cape**	363 NC061: Richtersveld	B3	3,1%	11 982	107
6	**Northern Cape**	371 NC073: Emthanjeni	B3	3,3%	42 356	820
7	**Northern Cape**	379 NC083: //Khara Hais	B2	3,7%	93 494	1094
8	**Northern Cape**	377+378: Mier + Kai Garieb	B3	3,8%	72872	853
1	**Northern Cape**	367 NC066: Karoo Hoogland	B3	4,2%	12 588	112
2	**Northern Cape**	365+368: Kamiesberg+ Khai-Ma	B3	4,7%	22652	202
3	**Northern Cape**	375 NC077: Siyathemba	B3	5,6%	21 591	418
4	**Northern Cape**	383 NC091: Sol Plaatjie	B1	5,6%	248 041	2349
5	**Northern Cape**	373 NC075: Renosterberg	B3	6,0%	10 978	213
6	**Northern Cape**	372 NC074: Kareeberg	B3	6,0%	11 673	226
7	**Northern Cape**	381 NC085: Tsantsabane	B3	6,5%	35 093	411
8	**Northern Cape**	369 NC071: Ubuntu	B3	6,9%	18 601	360
1	**Northern Cape**	385 NC093: Magareng	B3	7,6%	24 204	229
2	**Northern Cape**	361 NC452: Ga-Segonyana	B3	8,9%	93 651	1052
3	**Northern Cape**	370 NC072: Umsobomvu	B3	9,2%	28 376	550
4	**Northern Cape**	376 NC078: Siyancuma	B3	9,8%	37 076	718
5	**Northern Cape**	386 NC094: Phokwane	B3	10,0%	63 000	597
6	**Northern Cape**	384 NC092: Dikgatlong	B3	11,0%	46 841	444
7	**Northern Cape**	374 NC076: Thembelihle	B3	11,7%	15 701	304
8	**Northern Cape**	380 NC084: !Kheis	B3	12,3%	16 637	195
9	**Northern Cape**	360 NC451: Joe Morolong	B4	18,2%	89 530	1006
1	**Western Cape**	162 WC013: Bergrivier	B3	1,0%	61 897	522
2	**Western Cape**	164 WC015: Swartland	B3	1,0%	113 762	959
3	**Western Cape**	175 WC042: Hessequa	B3	1,5%	52 642	459
4	**Western Cape**	169 WC026: Langeberg	B3	1,7%	97 724	786
5	**Western Cape**	Cape Town Tygerberg Health Sub-District		1,74%	597282	4505
6	**Western Cape**	165 WC022: Witzenberg	B3	1,8%	115 946	933
7	**Western Cape**	Cape Town Northern Health Sub-District		1,88%	372579	2810
8	**Western Cape**	181+182:laings Prince albert	B3	2,1%	21425	209
9	**Western Cape**	173 WC033: Cape Agulhas	B3	2,1%	33 038	271
10	**Western Cape**	166 WC023: Drakenstein	B1	2,1%	251 262	2021

Rank of sub-districts by poverty headcount in each Tertile	Province	Sub-district	Municipal sub-category	2011 poverty headcount	2011 Population	Estimated 2013 deaths
1	**Western Cape**	163 WC014: Saldanha Bay	B2	2,2%	99 193	836
2	**Western Cape**	174 WC041: Kannaland	B3	2,5%	24 767	216
3	**Western Cape**	170 WC034: Swellendam	B3	2,5%	35 916	295
4	**Western Cape**	183 WC053: Beaufort West	B3	2,5%	49 586	484
5	**Western Cape**	Cape Town Southern Health Sub-District		2,72%	564609	4259
6	**Western Cape**	161 WC012: Cederberg	B3	2,8%	49 768	419
7	**Western Cape**	168 WC025: Breede Valley	B2	2,8%	166 825	1342
8	**Western Cape**	Cape Town Eastern Health Sub-District		2,89%	507153	3825
9	**Western Cape**	176 WC043: Mossel Bay	B2	3,2%	89 430	780
10	**Western Cape**	177 WC044: George	B1	3,3%	193 672	1690
1	**Western Cape**	160 WC011: Matzikama	B3	3,4%	67 147	566
2	**Western Cape**	Cape Town Western Health Sub-District		3,42%	470883	3552
3	**Western Cape**	171 WC031: Theewaterskloof	B3	3,7%	108 790	893
4	**Western Cape**	167 WC024: Stellenbosch	B1	3,8%	155 733	1253
5	**Western Cape**	178 WC045: Oudtshoorn	B2	3,9%	95 933	837
6	**Western Cape**	Cape Town Klipfontein Health Sub-District		4,15%	381480	2877
7	**Western Cape**	172 WC032: Overstrand	B2	4,6%	80 432	660
8	**Western Cape**	Cape Town Mitchells Plain Health Sub-District		4,64%	420924	3175
9	**Western Cape**	180 WC048: Knysna	B2	6,2%	68 659	599
10	**Western Cape**	179 WC047: Bitou	B3	6,3%	49 162	429
11	**Western Cape**	Cape Town Khayelitsha Health Sub-District		9,18%	425139	3207

8.3 Demographics of 27 sampled health sub-districts

Province	Code	Subdistrict	Category	2011 population	2013 deaths (N)	Poverty headcount (%)	Urban (%)	Formal housing (%)	African (%)	Coloured (%)	Asian (%)	White (%)	Other (%)
Eastern Cape	NMAC	Nelson Mandela Bay C	METRO	458364	2784	3.45	100	86.7	31.9	40.8	2.2	24.0	1.2
Eastern Cape	BUF	Buffalo City	METRO	755200	8043	9.30	82.3	72.5	85.1	6	0.8	7.7	0.3
Eastern Cape	EC154	Port St Johns	LOCAL	156136	1479	28.20	0.7	24.6	99.3	0.4	0.1	0.2	0
Free State	FS162	Kopanong	LOCAL	49171	885	3.80	84.3	91.4	71.5	18.2	0.4	9.4	0.5
Free State	FS192	Dihlabeng	LOCAL	128704	1612	6.10	82.6	76.7	87.4	1.5	0.5	10.4	0.2
Free State	FS194	Maluti a Phofung	LOCAL	335784	4205	7.90	39.2	80.4	98.2	0.2	0.2	1.3	0.1
Gauteng	JHBF	Johannesburg F Health Sub-District	METRO	608556	4044	1.03	100	92.8	64.2	13.9	6.7	13.9	1.3
Gauteng	EKUN2	Ekurhuleni North 2	METRO	541419	4437	5.51	99.9	73.7	52.3	8.0	6.8	32.1	0.8
Gauteng	EKUE2	Ekurhuleni East 2	METRO	333354	2732	7.14	99.9	79.0	80.8	2.3	1.9	14.6	0.5
KwaZulu-Natal	KZN232	Emnambithi/Ladysmith	LOCAL	237437	2372	8.10	53.1	80.5	91.8	1	4.4	2.7	0.2
KwaZulu-Natal	KZN227	Richmond	LOCAL	65793	679	13.00	27.6	54.7	95.2	0.9	1.1	2.6	0.2
KwaZulu-Natal	KZN272	Jozini	LOCAL	186502	1298	22.30	4.9	76.5	99.2	0.1	0.2	0.3	0.2
Limpopo	LIM335	Maruleng	LOCAL	94857	784	18.2	3.3	96	95.5	0.3	0.2	3.8	0.2
Limpopo	LIM342	Mutale	LOCAL	91870	668	22.4	2.9	87.3	99.3	0.1	0.1	0.5	0.1
Limpopo	LIM361	Thabazimbi	LOCAL	85234	658	25.6	82.2	70.7	84.3	0.6	0.2	14.4	0.4
Mpumalanga	MP305	Lekwa	LOCAL	115662	1047	4.5	81.9	73.6	84.2	2.9	1.2	11.4	0.3
Mpumalanga	MP312	Emalahleni	LOCAL	395466	3093	8.0	95.4	77.2	81.3	1.7	0.9	15.7	0.4
Mpumalanga	MP302	Msukaligwa	LOCAL	149377	1352	9.2	80.1	75.3	88.1	0.6	1.1	9.8	0.3
North West	NW403	City of Matlosana	LOCAL	398676	4443	4.6	93	82.9	81	3.5	0.8	14.5	0.3
North West	NW375	Moses Kotane	LOCAL	242554	2033	8.3	7.4	78.3	98.3	0.3	0.5	0.8	0.2
North West	NW381	Ratlou	LOCAL	107339	1163	16.6	0	89.9	98.2	0.7	0.2	0.7	0.1

Province	Code	Subdistrict	Category	2011 population	2013 deaths (N)	Poverty headcount (%)	Urban (%)	Formal housing (%)	African (%)	Coloured (%)	Asian (%)	White (%)	Other (%)
Northern Cape	NC083	//Khara Hais	LOCAL	93494	1094	3.7	94.7	75.2	23.1	65.2	0.7	9.9	1.2
Northern Cape	NC074	Kareeberg	LOCAL	11673	226	6.0	81.9	89.6	4.8	85.1	0.5	9.1	0.5
Northern Cape	NC451	Joe Morolong	LOCAL	89530	1006	18.2	2.4	72.5	96.4	2	0.3	1.2	0.2
Western Cape	WC013	Bergrivier	LOCAL	61897	522	1.0	60.3	93.4	11.3	70.9	0.4	16.9	0.4
Western Cape	WC041	Kannaland	LOCAL	24767	216	2.5	68.2	96.3	4.7	84.6	0.3	9.9	0.5
Western Cape	WC047	Bitou	LOCAL	49162	429	6.3	89.7	72.2	45.2	31.2	0.5	16.9	6.1
Total Sample				**5867978**	**53300**		**64.6**	**78.5**	**75.7**	**9.6**	**12.0**	**2.2**	**0.6**
Census 2011							**62.0**	**77.6**	**79.6**	**9.0**	**8.9**	**2.5**	

NOTES

Contact

Tel: +27 21 938 0911
Fax: +27 21 938 0200
Email: info@mrc.ac.za
www.samrc.ac.za