ASSESSMENT OF THE QUALITY OF ACUTE FLACCID PARALYSIS SURVEILLANCE DATA IN THE WORLD HEALTH ORGANIZATION AFRICAN REGION

KEITH SHABA

A mini-thesis submitted in partial fulfillment of the requirements for the degree of Masters in Public Health at the School of Public Health, University of the Western Cape

Supervisor: Professor Ehimario Igumbor

Co-Supervisor: Ms Verona Mathews

SEPTEMBER, 2012

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

List of Tables **Page number**

List of Figures

LIST OF ABBREVIATIONS AND ACRONYMS

AFP	Acute Flaccid Paralysis
BCG	Bacille Calmette-Guérin
CIF	Case Investigation Form
CDC	Centers for Disease Control and Prevention
CNS	Central Nervous System
CSF	Cerebral Spinal Fluid
DPT	Diphtheria, Pertussis and Tetanus
DQA	Data Quality Audit
DQS	Data Quality Survey
EPI	Expanded Programme of Immunization
EPID	Epidemiological Indication
GAVI	Global Alliance for Vaccines and Immunization
GBS	Guillain-Barre Syndrome
GPEI	Global Polio Eradication Initiative
Hib	*Haemophilus influenza* type b vaccine
HepB	Hepatitis B vaccine
mOPV	monovalent Oral Polio Vaccine
IPV	Inactivated Polio Vaccine
IST	Inter-Country Support Team
ITD	Intra-typic Differentiation
IVD	Immunization and Vaccine Development
JICA	Japan International Cooperation Agency

OPV	Oral Polio Vaccine
PAHO	Pan American Health Organization
STOP	Stop Transmission Of Polio
USAID	United States Agency for International Development
UWC	University of the Western Cape
UN	United Nations
UNICEF	United Nation Children's Fund
WHO	World Health Organization

DECLARATION

I declare that "*Assessment of the Quality of Acute Flaccid Paralysis Surveillance Data in the World Health Organization, African Region*" is my own work, that it has not been submitted before for any degree or examination in any other University or College, and that all sources I have used or quoted have been indicated and acknowledged as complete references.

Keith Shaba SEPETEMBER 2012

Signed:……

ACKNOWLEDGEMENT

The completion of this work could not have been possible without the guidance of my supervisors, Professor Ehimario Igumbor and Ms Verona Mathews of the School of Public Health (SOPH) at the University of the Western Cape and encouragement that led to the improvement of this document, and for this, am indebted to them. I would like to pay special thanks to Professor Bartholomew Akanmori for his support and dedication. I wish also to acknowledge Dr Levon Arevshatian, retired WHO Representative for his guidance and input. I wish to recognize, Dr Francis Kasolo, Dr Jason Mwenda, Dr Sam Okiror and Dr Toh, for their input at different stages of preparation of this document. Special thanks, also goes to Ms Corrine Carolissen and Ms Janine Kader, the School administrators for providing the necessary materials and latest information on the course. I would also like to thank my program manager, Dr Deo Nshimirimana for his moral support and allowing me to use the available data within the house for this study. Last but not the least I would like to thank all other colleagues, whose names I have not mentioned in this document, but contributed in one way or the other to its successful completion

ABSTRACT

Assessment of the Quality of Acute Flaccid Paralysis Surveillance Data in the World Health Organization, African Region

Keith Shaba

MPH, Mini-thesis, School of Public Health, University of the Western Cape

Background: Poliomyelitis (polio) is an infectious disease of high public health importance. In 1988, the World Health Organization (WHO) set the goal of polio eradication worldwide through the Global Polio Eradication Initiative (GPEI). A three-year period of zero indigenous wild poliovirus in all countries, in the presence of high-quality acute flaccid paralysis (AFP) surveillance, is the basis of an independent commission's determination of when a WHO region or a country can be certified as polio free. AFP surveillance being one of the critical elements in polio eradication campaign, aims to report and investigate all cases of acute flaccid paralysis occurring in children aged less than 15 years using clinical, epidemiological and laboratory methods. The information collected is cleaned and entered, into a database and maintained in EPI Info format at the WHO country office of each of the 46 countries, the three sub regional offices or Inter country Support Teams (IST) offices and the WHO African Regional Office. In addition, data from sixteen polio laboratories in various African countries maintain records of the laboratory findings and results of confirmed polio cases.

The quality of data generated through AFP surveillance and maintained in the African regional data base has not been critically and systematically reviewed and documented.

This study therefore was designed to gather information and document the quality of AFP data base, a key component of the global polio eradication effort.

Study Design and Sample Size: A cross-sectional descriptive study involving the retrospective review of clinical and laboratory databases of AFP surveillance over a five-year period (2004 - 2008) was designed. In this study, databases of CIFs containing clinical and laboratory data from AFP cases reported from all 46 countries of the WHO African Region comprising of 57,619 clinical and 59,843 laboratory records were critically reviewed.

Data Collection and Analysis: Databases of reported CIF covering the period 2004 - 2008 maintained in the Immunization and Vaccine Development (IVD) programme at the WHO Regional Office for Africa were accessed and checked for completeness and accuracy. Data was analyzed using EPI Info version 3.3.2. For the purpose of this study, an error framework was developed with 4 major codes A, B, C, D and 11 sub-codes reflecting different elements of AFP surveillance. The proportions of missing data and incorrect entries were calculated for clinical and laboratory databases and displayed using tables and graphs. Comparisons were made in the numbers of errors between various countries and sub-regions and proportions in the tables generated were used to assess the association between the level of accuracy and completeness of data.

Results: The study identified errors against all of 11 error codes in both the clinical and laboratory databases. Comparisons of data quality indicators between the clinical and

laboratory databases for the above period revealed that there are more errors in the clinical database than in the laboratory database.

Conclusions:

There are more errors in the clinical database than in the laboratory database. The analysis by major error groups in clinical database showed that the proportion of records with errors in groups A and B are not high (0.1-1.4%) compared with the other two groups (C and D). In group C, the dominant error group was C2 related to optimal time of specimen collection contributing 13.3% of all errors in 2005, reduced to 8.6% in 2008. The errors identified have implications for the Global Polio Eradication Initiative with some of them being critical for the documentation and processes required for the certification of the African Region as polio free.

Recommendations: Careful and frequent review of data and identification of errors and swift correction of these errors together with training and motivation of staff involved in data entry, collation, analysis and reporting as well as regular meetings to harmonize data should be pursued by all the countries.

KEYWORDS

Global Polio Eradication Initiative

Acute Flaccid Paralysis Surveillance

Clinical and laboratory databases

Data quality

Data errors

Completeness

Timeliness

Indicators

Laboratory Investigation and confirmation

Dates of onset of paralysis,

Date specimen collection

CHAPTER ONE

INTRODUCTION

1.1 Background

Overview of Polio Eradication initiative

Poliomyelitis is an important vaccine-preventable infectious disease which affects mostly children and causes permanent paralysis in those who survive, resulting in a major economic burden to the families and individuals. In 2009 there were 1,606 cases of wild poliovirus (WPV) infections worldwide, 388 of which occurred in Nigeria. In addition outbreaks of polio affected 19 previously polio-free African countries (Centers for Disease Control and Prevention 2010). The disease is thus an important public health problem and a major priority for eradication globally by the World Health Organization (WHO). The strategies for the eradication of poliomyelitis are: ensuring high routine immunization coverage (greater than 80%) with at least 3 doses of oral poliovirus vaccine (OPV), supplementary immunization conducted as national poliomyelitis immunization days (NIDs), implementation of surveillance to identify any cases of Acute Flaccid Paralysis (AFP) for detecting new cases of the disease and mop-up vaccination campaigns once poliomyelitis is reduced to focal (WHO, 2000). The strategies for polio eradication thus depend largely on the availability of high quality AFP surveillance and immunization coverage data.

Importance of Surveillance Information

The AFP surveillance data is used to determine which areas have transmission of the virus which causes poliomyelitis and together with the final laboratory results of virus isolation, identification and characterization allows the determination type of the polio vaccines (oral polio types 1, 2, 3, bivalent or trivalent vaccine) to use in response to stopping the spread of the virus in the population. In addition, by identifying the virus responsible for an AFP case, the origin of the virus can also be determined, in relation to viruses isolated from other regions or parts of a country or region, enabling response to be monitored in such areas as well. By serving as a measure of the success of the polio eradication programme, AFP data is also used for resource mobilisation in support of the Polio Eradication Initiative (PEI). For example, the GAVI Alliance has based its funding of immunization activities in developing countries on accurate, timely and easily verifiable immunization data, including AFP surveillance data (Lim Stein, Charrow and Murray, 2008). In some cases absence or inaccurate immunization data has led to the suspension of GAVI Alliance funding which in turn has adversely affected the PEI. Thus as routine immunization coverage improves the number of cases of disease decreases and surveillance is required to identify any cases occurring in those not immunized and to target them with supplementary immunization so as to eradicate or eliminate or control the disease eventually. This is critical for polio which has been targeted for eradication requiring that transmission is completely stopped.

AFP Surveillance

Acute Flaccid Paralysis surveillance is a systemic monitoring for the presence of paralysis with a sudden onset, which could indicate poliomyelitis. The identification of cases of AFP requires laboratory investigation of stool specimens for presence of the polio virus. The information generated is used for the GPEI. There are three main ways of conducting AFP surveillance (WHO, 2008):

i) Active surveillance: is defined as regular visits by surveillance officers to health care facilities (clinics, hospitals, rehabilitation centres and traditional healers' premises) to search for and investigate unreported cases of acute flaccid paralysis through a review of health records.

ii) Passive Surveillance: this involves health facilities sending a report of AFP cases identified in the facility during the reporting period.

iii) Surveillance by community focal points: community focal persons (traditional healers, birth attendants, chiefs) who are likely to know about the occurrence of AFP case in the community are orientated on AFP symptoms so that they can report these cases to surveillance officers or health centres.

Case Investigation Form

All the information about each AFP case is carefully collected using a case investigation form (CIF) - Annexure 1. The CIF is used to record case information which comprises of patient information, demography, clinical and laboratory information, and follow up information. The CIF is entered into two separate databases, one for the clinical data and the other the laboratory data. The clinical database contains all information obtained after a clinical examination of the

AFP case by a health worker, while the laboratory database contains all the results of laboratory investigations of the stool specimen collected from the AFP case. These databases form the AFP Surveillance system.

Clinical and Laboratory Database

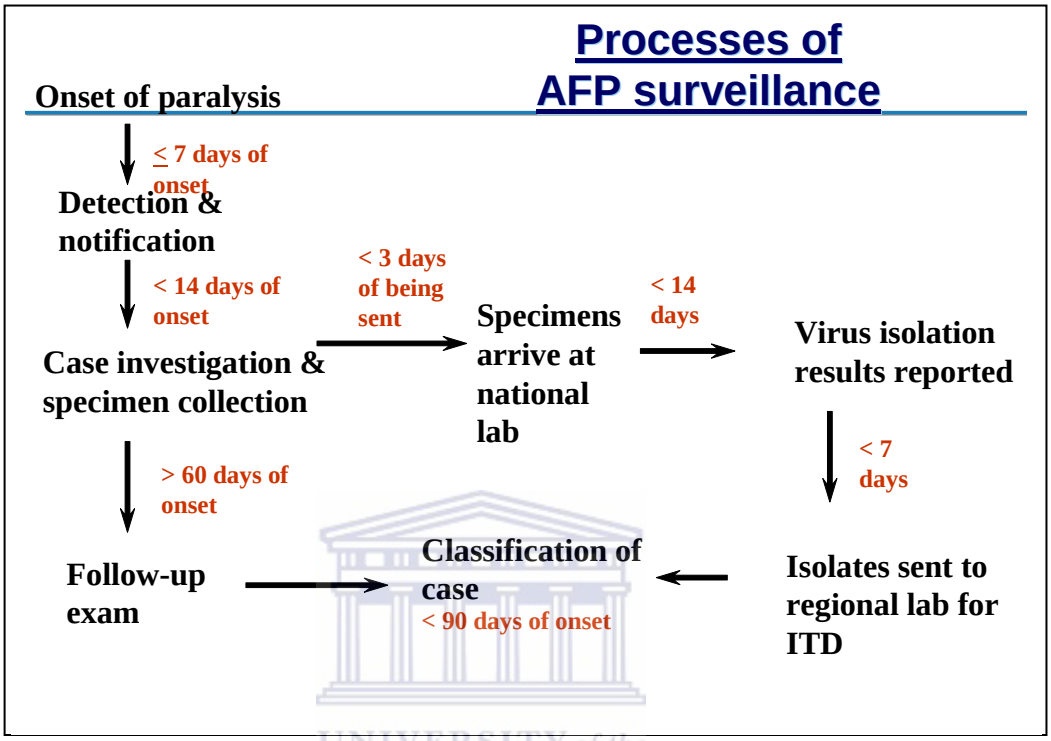
An AFP clinical database is a collection of surveillance information derived from a case investigation form filled in at field level for acute flaccid paralysis. A laboratory database is a collection of information related to laboratory diagnosis and confirmation of the AFP.

For the AFP surveillance system to be effectively implemented there is need for harmonisation of the two datasets or databases, in order to ensure that all information from the CIF is accurately entered in both datasets (WHO, Surveillance Guidelines, 2005).

The process of AFP surveillance: The clinical databases are generated at the country level, and transmitted to their respective Inter-country Support offices (ISTs). The ISTs merge and clean the databases before onward transmission to the WHO Regional office (AFRO) based in Brazzaville. On the other hand, the laboratory databases originating from country level with clinical information filled and the laboratory section updated in the laboratory, are sent directly from the laboratories to the WHO Regional office where merging and cleaning are done.

The people responsible for the cleaning and merging of the two databases are the surveillance officers, epidemiologists and data managers. The person that is overall in charge of the AFP Surveillance is the WHO Regional Epidemiologist.

Figure 1 Diagram illustrating the steps followed during AFP surveillance with accompanying stipulated timelines



Exam should read "clinical examination"; lab should read "laboratory"; ITD should read "intra-typic differentiation"
Source: WHO, Surveillance Guidelines, 2005.

The processes involved in AFP surveillance as well as the timelines are illustrated in figure 1. The process begins by identification of a suspected case of AFP defined as a child under 15 years of age with acute (sudden onset) flaccid paralysis (weakness of the limb, arm, leg or both) or any person of any age when paralytic illness or polio is suspected by a clinician (Marx Glass and Sutter, 2000).

Once the suspected case has been reported, a surveillance officer visits the suspected case to confirm that it is a true case of AFP. This health worker interviews the parents or guardians of

the case and conducts a clinical examination of the case before completing the case investigation form (CIF) and obtaining two stool specimens. The stool specimens collected are sent with a copy of the CIF to the designated laboratory for virological testing. The information generated from the laboratory is sent back to the surveillance officer who then notifies the health centre on whether the case is confirmed as polio or not. All the processes have specified timelines within which they should be carried out as shown in the figure and which also serve as the basis of the indicators for monitoring the programme.

When the case is confirmed as poliomyelitis the clinical information on the CIF is entered into the clinical database while laboratory information is captured in the laboratory database. Errors can occur at different levels such as when the CIF is being filled and when information is being entered in the clinical and laboratory databases. In addition, the entry of laboratory results at program level could also be an avenue for introducing errors. Delays can also occur at any of the stages, affecting the timeliness of reporting.

Once the clinical and laboratory data are received at the IST and WHO regional offices, feedback is provided to the source of information in the form of a line listing highlighting all areas of concern. Where data gaps have been identified, possible solutions are suggested to correct errors and these are monitored to ensure that they have been corrected in the subsequent databases received. Before feedback is provided, the data quality checks would have been performed beginning with the CIF and subsequently the data that has been merged.

For this particular study, a standard template categorizing the types of errors in the two inter related databases, was developed in order to highlight the differences in the two sources of data. The categories chosen focused mainly on the five core variables that are used to calculate AFP surveillance indicators and monitor program performance (refer to annex)

1.2 Problem Statement

It has been found that the capture of CIF data for AFP surveillance has often resulted in numerous errors. The consequences of errors in the database and delays in rectifying them are that cases of poliomyelitis are overlooked or responses in the form of supplementary immunization or mop-up campaigns could be delayed, leading to the spread of the disease and delays in eradication efforts. Despite remedial actions such as regular follow up, monthly data harmonization meetings, to rectify errors arising from data entry as well the publication of feedback bulletins focussing on data discrepancies, the quality of AFP data has not significantly improved and remains a major challenge to the polio eradication programme in Africa (WHO, 2005). Significantly, no systematic identification, quantification and categorization of the types of errors in the AFP surveillance data have been carried out in recent years.

1.3 The Purpose of the Study

The purpose of this study was to identify the types and to quantify the extent of errors in reported clinical and laboratory AFP datasets. Also to make recommendations aimed at improving data quality for AFP surveillance in the WHO Africa Region and to contribute to the subsequent eradication of polio from the continent.

1.4 Study Objectives

1.4.1 Aim

To assess the quality of clinical and laboratory data for AFP surveillance in the WHO Africa Region.

1.4.2 Specific Objectives

1. To identify the types of errors in the variables entered in the clinical and laboratory databases for AFP surveillance.
2. To quantify the types of errors identified in the databases for various countries in the WHO African Region.
3. To compare the errors, completeness and timeliness of reports in the clinical and laboratory databases for different countries and sub-regions.

CHAPTER TWO

LITERATURE REVIEW

2.1 Poliomyelitis as a Disease

Polio is an enteric virus and expresses three important antigens (type 1, 2 and 3) associated with paralysis. These antigens are used to classify the three serotypes and are also the targets for vaccine development. Poliomyelitis also referred to as "polio" is a highly contagious, acute viral disease (Murray, Lopez and Mathers 2004). The virus is transmitted from person to person through the feaco-oral route with occasional transmission via milk, sewage and drinking of contaminated water (Pavlov *et al.* 2005).

Though polio mainly affects children below 3 years old, it can also affect older children and adults. The disease progression was not clearly understood until 1870 that the lesions associated with poliomyelitis in the spinal cord were first described (Sutter, Kew and Cochi, 1988). More recent research on poliomyelitis as a disease has focused on transmission, morbidity, and the disease burden. One in every 200 affected children becomes paralyzed, followed by permanent disability (Bernier, 1984, WHO, 2005), resulting in huge economic burden for the family and the country. The high mortality rate and the disabilities attributable to polio infection have increased the economic burden especially at the household and community levels (Thompson *et al.* 2008) About 5%- 10% of affected children with paralytic poliomyelitis have a fatal outcome (Kenji Shibuye, Murray, 1984).

For an estimated global birth cohort of 125 million for 2004, an estimated 600,000 people developed permanent disability due to paralytic poliomyelitis and up to 60,000 of these died from the disease. An additional burden of the disease is the high cost of rehabilitation and physiotherapy for survivors. More resources are needed to provide support especially to disabled children (Murray, Lopez and Mathers 2004; Biritwum *et al.* 2001).

The World Health Organization (WHO) and partners are collaborating to eradicate poliomyelitis worldwide. To monitor progress, countries perform surveillance for acute flaccid paralysis. The WHO African Regional Office (WHO-AFRO) and the U.S Centers for Disease Control and Prevention (CDC) are also involved in strengthening communicable disease surveillance and response in Africa (http://www.biomedcentral.com/1471-2458/2/27/, Accessed on 15/05/2010).

2.2 Clinical Manifestation of Poliomyelitis

Infection with poliomyelitis is most often recognized by the acute onset of flaccid paralysis. Poliovirus infection occurs in the gastrointestinal tract and spreads mainly to the lymph nodes and central nervous system. Flaccid paralysis occurs in less than 1% of poliovirus infections; over 90% of infections are either unapparent or result in a non-specific fever (WHO, 2005).

The paralysis of poliomyelitis is usually asymmetrical with fever present at onset. The maximum extent of paralysis is reached in a short period of time, usually 3-4 days. The legs are more affected than the arms. Paralysis of muscles supporting breathing can be life-threatening. Some improvement in paralysis may occur during convalescence but paralysis persisting after 60 days is likely to be permanent.

In paralytic polio, about 50% of patients recover with no residual paralysis, with about 25% left with mild disabilities, and the remaining 25% have severe permanent disabilities such as impaired hearing, speech, hydrocephalus and kyphosis (Lund *et al.* 2007; Lund and Lexel 2008; Lund and Nygard 2004).

2.3 Prevention and Control of Poliomyelitis

There is no treatment for poliomyelitis. Vaccination is the only cost effective method for the prevention and control of poliomyelitis (WHO Fact sheets, 2009). Equally important is the monitoring of protective immunity and maintaining surveillance to sustain the present polio-free situation (Nokleby *et al.* 2010). Subsequently several landmarks have been achieved including the development of effective vaccines and introduction of the goal of global eradication of the disease by 2005 (Murray, Lopez and Mathers 2004; WHO, 2000). Through vaccination most of the developed countries have eradicated poliomyelitis. The Global Poliomyelitis Eradication Initiative and the GAVI Alliance are providing funding for the eradication of this disease in African countries and within the Expanded Program on Immunization (EPI), which targets children under-five years old. The programme provides free of charge, the oral poliomyelitis vaccine (OPV) during routine immunization activities at health facilities and during national immunization days (NIDs) through fixed, door to door and outreach strategies.

There are two main types of polio vaccines developed by American scientists Salk and Sabin available: the more expensive inactivated polio vaccine (IPV), which is administered by injection

and the much cheaper, live oral polio vaccine (OPV) given orally (Murray, Lopez and Mathers 2004).

In 1963, trivalent live oral polio vaccine (Sabin), incorporating the three types of poliomyelitis viruses (1, 2, 3) replaced IPV as the recommended polio vaccine. Most countries have commonly used oral polio vaccine for preventing poliomyelitis because of the ease of administration and enhanced mucosal immunity, low cost and the ability to boost population immunity and subsequently interrupt transmission of the virus. Secondary spread of this live Sabin-derived vaccine virus can occur from vaccinated to unvaccinated close contacts conferring herd population immunity (WHO Weekly Epidemiologic Record, 2003).

The development and licensure of monovalent types 1 and 3 live oral polio vaccines (mOPV1 and mOPV3) have provided the Polio Eradication Initiative with new and additional vaccines against disease caused by the specific poliovirus types 1 and 3 that are prevalent in Africa. For example, the use of mOPV1 in Nigeria in 2006 was partly responsible for the remarkable reduction in the wild poliovirus type 1 infections experienced in 2007 (WHO, 2008).

2.4 Global Polio Eradication Initiative

The Global Polio Eradication Initiative was launched at the 41st World Health Assembly in 1988, with the aim of eradicating poliomyelitis by 2005 (WER No. 36, 2009). The initiative involves several partners such as Rotary International, USAID, CDC, JICA, UNICEF and many others. Due to persistence of wild poliovirus in Asia and African countries, the 2005 target was not achieved. Supported by the World Health Organization and the Center for Disease Control and

Prevention, the STOP (Stop Transmission Of Polio) initiative was later created with the aim to eradicate Polio from the globe by the Year 2010 (Nsubuga *et al*. 2002).

The Global Polio Eradication Initiative has made remarkable progress showing the incidence of polio has decreased from an estimated 350,000 cases annually to 1,655 reported over the period 1988 to 2008 and 1606 WPV cases in 2009 (WHO, 2008, WER No.12, 2009, WER No.18, 2010). Global vaccination efforts have contributed to this progress. By 2005, indigenous transmission of wild poliovirus (WPV) types 1 and 3 (WPV1 and WPV3) had been eliminated from all but four countries (Afghanistan, India, Nigeria, and Pakistan). No cases of WPV type 2 have been reported since 1999. Despite the progress, Nigeria and India still remain polio-endemic countries and some sporadic WPV cases have been notified in some other parts of Africa and Asia. Early this year new cases of paralytic polio due to WPV have been reported in Tajikistan raising fears of the re-emergence of polio in parts of Europe several decades after its eradication from the region (WHO 2010).

2.5 Polio Eradication in the African Region

The number of polio-endemic countries in the African Region has declined from 11 in 2000 to just one, Nigeria in 2008, with an 89% decrease in the number of reported cases of WPV within the same period (Polio News, 2009). This success has been attributed to high immunization coverage during mass vaccination campaigns, good AFP surveillance, appropriate logistics and the availability of funds. Currently there are 15 other countries in Africa with wild poliovirus, largely due to importations of the virus from Nigeria and so constituting a major setback to the eradication initiative on the continent (WER No.14, 2009, WER No16, 2009). The last few cases

were reported from Kenya and Uganda signaling that the threat of polio is still great for the African continent. Currently WPV is transmitted in the Horn of African region (Ethiopia, Kenya and Uganda) and in West Africa (Togo, Nigeria, Cote d'Ivoire, Burkina Faso, Mali and Niger) (see figure 2 in annex).

2.6 Surveillance for AFP and Laboratory Identification of Polioviruses from Stool Samples

Remarkable progress towards polio eradication has been noted globally (WER, 2009). Surveillance for cases of AFP and examination of stool specimens from AFP cases for the detection of poliovirus provides important data to target supplemental immunization activities (WER 1998). This is critical for the planning of immunization campaigns using the appropriate OPV type to ensure that the virus is eliminated in line with the GPEI's requirement for the eventual certification of polio-free status of a country or region (Kew *et al.* 2005).

The presence of AFP as a proxy indicator of infection with poliovirus with a high sensitivity level has strengthened the surveillance system by capturing all cases of poliomyelitis including atypical presentations and identifying cases of poliomyelitis due to spread of the vaccine strains. The number of AFP cases notified per 100 000 children less than 15 years within a given time depict the quality of surveillance activities, even when there are no cases of poliomyelitis (WHO, 2008). A global network of accredited laboratories has been established for the processing of all stool samples in support of this surveillance.

Standards have been set for the certification of polio-free status. Among others the certification standards require that countries continue to demonstrate high levels of surveillance indicators (WHO, WER No.18, 2004).

2.7 Challenges in Monitoring AFP Surveillance

Surveillance for poliovirus remains the cornerstone of monitoring the progress of the poliomyelitis eradication initiative (PEI). The process includes notification and careful investigation of cases of acute flaccid paralysis (AFP), collection of stool specimens from the suspected patient in a timely manner, and virological examination of the specimens in a WHO accredited laboratory for the confirmation of polio. Successful high quality AFP surveillance requires adherence to the approved Standard Operating Procedures and smooth interplay of several different groups involved in obtaining surveillance data (Hovi, 2006).

The performance of AFP surveillance is based on a number of indicators as shown in Table 1 (WHO, 2003). The accuracy of each of these indicators is determined by the *timeliness* and *completeness* of reporting at each level of collection as well as proper handling of stool specimens, including conditions of their shipment, and the capacity of the laboratory to perform the required analysis.

There are several factors which affect the quality of AFP surveillance and these include data quality, availability of resources, and training of health personnel in surveillance including laboratory staff and community participation in surveillance at grassroots level. The present study will focus on the quality of data as a major affecting surveillance for AFP.

2.8 General Overview on Data Quality

In the context of disease control programs, ensuring high quality of data requires going through the data periodically, updating, standardizing and correcting duplicate records to create a single record of the data, even if it is stored in multiple disparate systems (WHO, 2001). Data quality encompasses ensuring the reliability and authenticity of data for use in effective decision-making for disease control and prevention.

There are many reasons as well as several ways in which the quality of data may be compromised. These include introduction of errors during data entry, errors in analysis and reporting due to differences in data collection and reporting formats, inadequate performance by personnel and high staff workload (Harris *et al.* 2003; de Gourville *et al.* 2006; Danovaro *et al.* 2007). In a comprehensive study in Mozambique, the quality of quarterly aggregate summary data compiled and reported for antiretroviral treatment facilities was compared with Ministry of Health data for the same period. The study revealed that many health facilities reported inaccurate data and several facility-level characteristics (Makombe *et al. 2008),* as already enumerated were associated with data quality. Some of the reasons stated may be responsible for the poor quality of immunization data, including AFP Surveillance data at national, country, the inter-country support, laboratories and the WHO Regional Office. However this requires confirmation through a careful and systematic review.

The improvement of the quality of data will require identification of the sources and levels of errors and the competence of all personnel involved in the handling and transmission of AFP data and corrective measures put in place (Birmingham *et al.* 1997).

2.9 Vertical versus Integrated Surveillance Systems

Typically each vertical health program runs its own system with little regard to how information is integrated or is beneficial to the overall system. Centralization of statistical systems is advantageous where technical resources and expertise are in short supply (Macfarlane, 2005). However, most developing countries operate fragmented and disjointed disease surveillance systems resulting in collection of unreliable and unrepresentative data. The insufficient number of staff and their low capacity has negatively impacted on the timeliness and completeness of AFP surveillance data in most developing countries especially in Africa. The current AFP data collection system is regarded as vertical involving collection of data from the periphery and transfer progressively to the central level on a weekly basis.

2.10 Country Experiences and Research in Quality of Disease Surveillance

In 1985, the Pan American Health Organization (PAHO) adopted an initiative to eradicate poliomyelitis from the Western Hemisphere. They conducted a study to screen cases of AFP for poliomyelitis eradication as a way of improving specificity. In a cross sectional retrospective study conducted in the countries of Latin America, all AFP affecting children less than 15 years old were reported and investigated. Data quality was a critical point at the stage of polio eradication indicating the absence of wild poliovirus circulation in the WHO American Region (Andrus *et al.* 1992). The extent to which AFP data differs for different countries in Africa is not very clear because few studies have been conducted.

The inaccuracy of data has been reported in other studies in Africa. For example in a study carried out in Malawi to assess the quality of data, the accuracy of data was classified as

inadequate if there was 5% or more missing variables. This study revealed that while most sites offering health services had complete case registration and outcome data, many sites did not report accurate data for several critical data fields such as reasons for starting the treatment, outcome of treatment, and the treatment regimen. The national summary using the site reports resulted in 12% less in total number of persons on first-line treatment (Makombe *et al.* 2008).

In another study conducted in Mozambique (Mavimbe, Braa and Bjune 2005), the quality of the measles reporting system was studied. There were significant differences in data between facility reports and tally cards. Similar differences were encountered in data from health facilities when compared to district health data. Routine practice during supervision visits failed to focus on data inconsistency between the tally sheets and the facility reports.

Assessing the routine surveillance data for measles Jagrati and others showed a significant under-notification of measles cases resulting in inappropriate evidence-based decisions and not allowing in-depth analysis to monitor measles epidemiology in the country (Jagrati *et al*., 2006). In the Americas, a study assessing the immunization coverage data and procedures used to calculate coverage figures using DPT3 showed clear incompleteness and inconsistencies in data analysis (Danovaro *et al*, 2007). The studies reviewed are some of the examples of inaccuracies in data from health programmes in different parts of the world.

2.11 Assessment of Data Quality

Disease surveillance is necessary to identify priority areas and plan strategies of operation. The surveillance system needs to be strengthened for the collection of data on EPI target diseases for

measuring the impact of immunization on reduction of morbidity, mortality, and disability (Basu 1982). In general, there are several ways to assess the quality of data. These include periodic reviews of the data, prospective studies to evaluate the whole process from capture, entry, transfer, analysis and reporting as well as use of continuous monitoring tools. The GPEI has specific tools which have been made available to countries to enable them assess data quality, identify errors and their sources and to correct them in a timely fashion. These methods include the Data Quality Survey (DQS) (Woodard *et al.* 2007), monthly data harmonization meetings, publication of feedback bulletins with focus on quality and programmatic issues, and the use of cluster surveys (WHO, 2009). Despite the availability and use of these tools by countries in Africa the quality of immunization data in general and AFP surveillance data in particular still has inaccuracies, year after year.

2.12 Challenges of Data Quality in the African Region

Data quality is a major concern in public health programs especially the inaccuracies in immunization data. The extent of inaccuracies in immunization data in the African Region compared to other regions has not been well established.

Available reviews of routinely collected health records suggest that mechanisms to ensure data quality through supervision, harmonization of data and collection tools, and data quality assessments contribute to the quality, timeliness and accuracy of data and thus correcting these factors would improve data management practices (WHO Weekly Polio Updates, 2008)

The Global Alliance for Vaccine and Immunization (GAVI) has developed a standard tool for the quick assessment of the quality of data, the Data Quality Audit (DQA) (Ronveaux *et al.* 2005). This audit involves the provision of a standard data checklist for countries to use in auditing their data but has never been used to formally study data quality in Africa (WHO, 2005). A simulation has however revealed that the DQA lacks the precision it is supposed to have due to inadequate sample size (Woodard *et al.* 2007). Despite this the DQA is recommended as an innovative tool for the independent assessment of the quality of immunization monitoring systems, by identifying inconsistency in immunization data and providing a means of improvement of data quality (Ronveaux *et al.* 2005).

There is clearly an urgent need to determine the quality of AFP surveillance data within the GPEI in the WHO African Region and to identify the factors which could compromise the quality of data and propose ways to improve data quality in the region.

This study was designed to use available proven tools to retrospectively review the AFP surveillance data in the African region.

CHAPTER THREE

METHODOLOGY

3.1 Study Design

This was a cross-sectional descriptive study involving the retrospective review of clinical and laboratory databases of AFP cases collected over a five-year period (2004 - 2008) by 46 Member States of the WHO African Region.

3.2 Study Area and Population

All the clinical and laboratory records including case investigation forms of AFP cases reported over a five-year period (2004 to 2008) from 46 countries involved in the GPEI program in the WHO Africa Region were used for the present study. These countries were grouped by WHO Regional Office for Africa (AFRO) for operational purposes into three sub-regions referred to as Inter-Country Support Teams (IST): West IST- with 17 countries, Central IST- with 11 countries and South East IST with 18 countries (Appendix 2).

3.3 Sample Size

A sample size was not estimated for this study as the entire sum of reported AFP cases from 2004 to 2008 (five years), comprising 57,619 clinical and 59,843 laboratory records were reviewed.

3.4 Sampling Procedures

No sampling procedures or techniques were required for the present study as it was based on the review and analysis of previously collected data. This required systematically going through the databases and identifying errors such as missing values, wrong entries and incomplete entries. A standard tool developed by the WHO (WHO, 2005) for the identification and logging of errors was used.

3.5 Data Collection

Database of reported CIF covering the five year period, 2004 to 2008 maintained in the Immunization and Vaccine Development (IVD) programme, at the WHO Regional Office for Africa was accessed and checked for completeness and accuracy. The databases by year of reporting are kept in EPI Info software on a shared drive at the IVD programme. Permission was sought from the IVD Programme Manager and all the administrative arrangements were made prior to the use of the data. No specific data collection techniques were required as previously collected data was used for the study.

3.6 Data Analysis

The different types of errors in the databases were identified and all tabulated. Comparisons of the total numbers and proportions of the different types of errors were compared for different variables such as sub-regions and countries, types of errors and trends over the past five years.

Data was analyzed using EPI Info version 3.3.2 (CDC, 2006). A report was considered complete if it contained critical variables (EPID number, date of onset, dates of 1st and 2nd stools

collection, date specimen received in the laboratory, primary and ITD laboratory results) used calculate indicators while accuracy was determined by correct matching between surveillance and laboratory databases. Proportions were compared using different tables produced (Figure 3) to determine the number of errors in each given category.

3.7 Validity and Reliability

Completeness and accuracy of data have been clearly defined to ensure that there is standard methodology of assessing the errors. A previously validated data quality assessment technique (WHO, 2005) was used to assess the quality of data while the surveillance guidelines protocol was used to assess the methodology of generating outputs. Selection bias has been eliminated as all records for the study period were reviewed. The reliability of assessment of errors was assessed by re-assessing a 1% sample of records for levels of error and completeness. Data was subjected to standard statistical analysis.

3.8 Generalizability of Results

The use of the entire population allows the results of the assessment of quality (accuracy, completeness and timeliness) to be generalizable to the situation in the 46 countries of the WHO Africa region for the study period. The outcome may also be applicable to other health related data collected by programmes in the African region since practices and policies for the various programmes of the Ministries of Health (MOHs) are similar and based on guidelines provided by WHO. The outcomes of the present study are however limited to the period covered by the data and cannot be applied to any other period (previous or later).

3.9 Limitations

While the present study assessed the quality of data and the common factors which could have compromised data quality, the study faced a number of limitations characteristic of similar studies. It was impossible to adduce all accurate reasons for the errors which were identified since the study was retrospective.

3.10 Ethical Statement

Ethical approval of the proposed study was sought from the Faculty of Community and Health Sciences and the UWC Ethics Committee. Permission was also sought from the Programme Manage, Immunization and Vaccines Development Programme of the Division for the Prevention and Control of Communicable Diseases, Regional Office for Africa, World Health Organization. Names and other information which could identify individuals were deleted to ensure the protection of the rights of individuals to privacy, which is consistent with guidelines of Ethical research involving human subjects (ICH Guidelines).

3.11 Data Storage and Cleaning

Data are stored electronically in EPI-Info format at WHO level and hard copies filed at country level. The data which was used in the present study had been cleaned during monthly data harmonization meetings, from feedback, and regular flagging of data quality issues in the modules used. Data cleaning programs components have been incorporated in all modules used for data capturing at country and WHO levels.

CHAPTER FOUR

RESULTS

The study consisted of the analysis of clinical and laboratory databases for 46 countries of the WHO African region, received at the WHO Regional office for the period 2004 – 2008.

4.1 Number of AFP Cases Expected versus Reported in Clinical Database

One of the key indicators of polio surveillance is the reported AFP rate. The standard indicator for the certification of a region as being free of poliomyelitis is pegged at 1 non-polio AFP case reported per 100,000 people below 15 years of age as an indication of a functional surveillance system (WHO Regional Office for Africa, 2005).

Table 1 below shows the expected AFP cases by year versus number of AFP cases reported using the above criteria.

Table 1: Expected AFP cases by year versus number of AFP cases reported in clinical database (1/100,000 population <15 yr.)

Year	Expected AFP cases *	Reported AFP cases	Reported AFP rate **
2004	2917	8560	2.9
2005	3189	11341	3.2
2006	2743	12067	3.8
2007	2873	11355	3.7
2008	2942	14296	4.5

* = Expected number of AFP cases is computed by dividing countries' population for under 15 years by 100,000. These are estimates.

** = Reported AFP rate is calculated using actual reported AFP cases for given year divided by under 15 year olds population

The table above indicates that throughout the years under study, the reported AFP cases far exceed the expected level indicating that countries' performance for this indicator is in general satisfactory.

In order to capture wild poliovirus among reported AFP cases, WHO established a standard of two stool specimens per single reported AFP case to be taken within 14 days of onset of paralysis, 24-48 hours apart. Table 3 below represents compliance of the countries in the African Region with this requirement.

Table 2: Total number of AFP cases reported versus number of stool specimens taken in the clinical database *)

Year	Total # AFP Cases reported	Stool samples taken		Specimen not taken (%)
		1 sample (%)	2 samples (%)	
2004	8560	69 (0.8)	8468 (99.1)	23 (0.3)
2005	11341	82 (0.7)	11230 (99.0)	29 (0.3)
2006	12067	14 (0.9)	11920 (98.8)	33 (0.3)
2007	11355	78 (0.7)	11255 (99.1)	22 (0.2)
2008	14296	87 (0.6)	14176 (99.2)	33 (0.2)
Total	57619	430 (0.7)	57619 (99.1)	140 (0.2)

*) Note: The clinical database is used to avoid duplication with the data in laboratory database. The clinical database is the initial step in AFP investigation leading to specimens being delivered to the laboratory.

The data in Table 2 indicates high compliance of countries (99.1%) with the standard requirement of collecting two samples per each reported AFP case. The proportion of AFP cases without a sample or with only one sample taken is low ranging from 0.2 to 0.3% and 0.6 to 0.9% respectively during the study period. This is not negligible, because as proportions of the total number of AFP cases, 140 potential cases of poliomyelitis could be missed leading to spread of the virus, polio being a highly infectious viral disease. When an AFP case is missing stool

specimen, such a case (after being notified) is normally followed up within 60 days to verify whether or not the patient eventually develops residual paralysis. If found to have residual paralysis or lost during follow up, (misplaced, death, etc.), such a case is classified as "Polio compatible" otherwise, it is discarded. However within the 60 days there is a likelihood of the transmission of the virus if indeed the AFP case turns out to be polio. This becomes a serious setback to the PEI program.

4.2 Status of clinical and laboratory databases

Proportion of annual AFP cases reported by year is shown in Figure 1 which indicates an increase in the number of reported AFP cases in the clinical database although throughout years under study a slight decline was observed in 2007 while the number of AFP cases in the laboratory database continued steadily to increase.

Figure 2: Reported AFP cases in clinical and laboratory databases, 2004-2008.

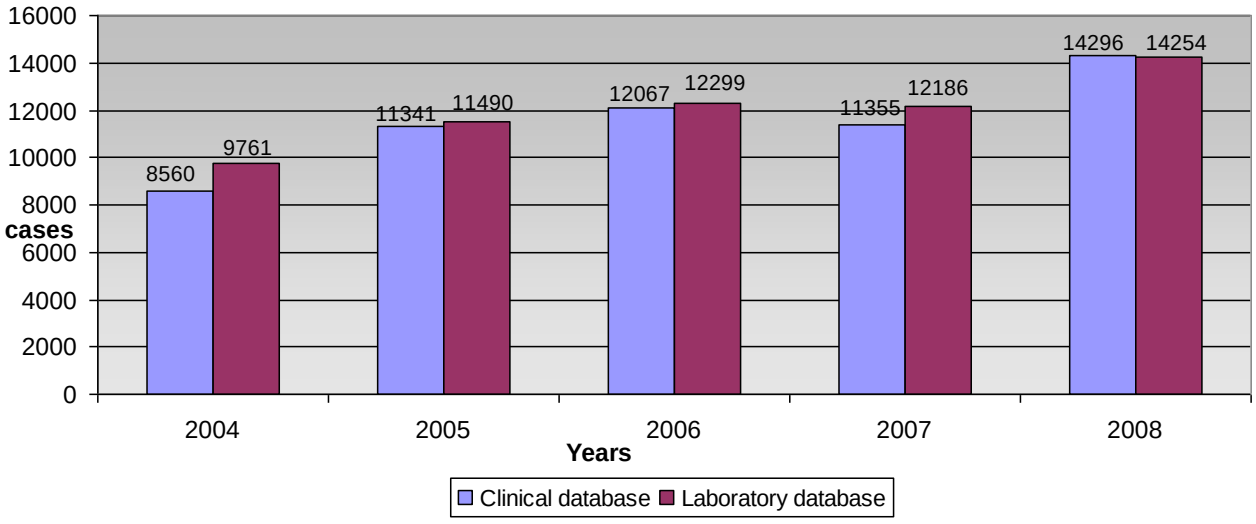


Figure 2 also provides data showing that the annual trend of reported AFP cases in the laboratory database was always more than in the clinical database with the exception of the year 2008. The

data shows a steady increase in the number of cases in both the clinical and laboratory databases from 2004 to 2008, except in 2007 when there were declines in numbers of cases for the two databases in 2007.

4.3 Identification of Errors in the clinical and laboratory databases

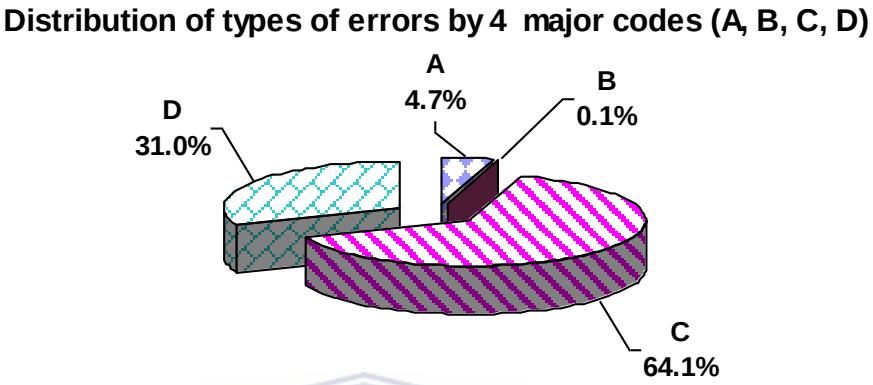
Data cleaning process using error identification program detailed in Appendix 2 revealed a total of 11 different types of errors distributed in 4 major groups (A, B, C, D) reflecting different elements of AFP surveillance: EPID number, date of onset, date on which stool sample was taken and the laboratory results as shown below in Table 3. This program was designed to identify gaps in data which can broadly be categorized as incomplete and wrongly formulated EPID numbers, missing information in key variables used for calculation of indicators of performance, illogical date entries leading to negative intervals and mismatches of information among related fields.

Table 3: Different types of errors identified in the clinical and laboratory databases, 2004-2008

Error type CODEs		Identified errors in the clinical and laboratory databases
A	A1	Wrong formulation of EPID Number (dash is missing in EPID number to separate codes)
	A2	Incomplete EPID Number (year or case number has an "x" or blank space)
	A3	Year in EPID Number and in the date of ONSET do not match
B	B1	Cases missing date of onset
	B2	Cases with year of onset not matching year of the database
C	C1	Date of onset and date 1st stool collected with negative interval
	C2	Cases where date of onset and stool collected is greater than 14 days
	C3	Cases where the interval between dates of 1st and 2nd stool collected is negative or >48 hours
D	D1	Cases with missing date specimen received in the laboratory
	D2	Cases with missing primary isolation results
	D3	Cases with missing intra-typic differentiation results

The analysis included a total of 17,857 error cases categorized in 4 major codes with their individual proportions shown in Figure 3.

Figure 3



Frequency of identified errors was highest (64.1%) for error group "C" related to stool specimen collection dates followed by group "D" related to laboratory results, which was 31.0%. The other two groups A (EPID number discrepancies) and B (inaccuracies related to date of onset) constituted small proportions among the total number of errors (4.7% and 0.1%). Description of errors per type of database is given in the following sections.

4.4 Errors in the Clinical Database

Analyses in this database were conducted on 57,619 records of AFP cases. Figure 4 represents the proportion of errors identified in this database among all AFP cases entered showing 20% level which by surveillance standards is considered to be high.

Figure 4

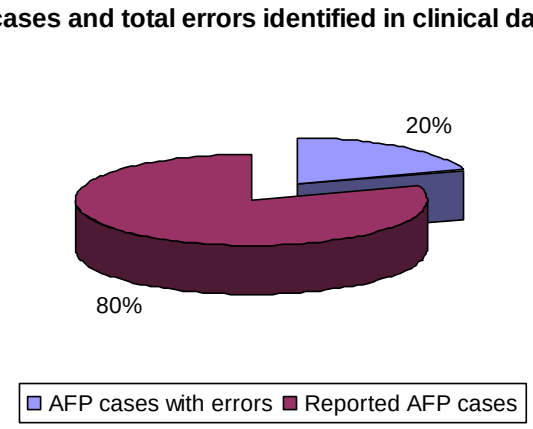


Table 4 below details the numbers and proportion of errors against 11 specific error codes identified for the period under study. Proportions of errors were calculated using their number over total reported cases while for proportions of missed ITD results – error code D3, the denominator was the number of positive isolates based on information given by the laboratory.

Table 4: Number of errors in the clinical database, African Region, 2004 - 2008

Type of error	2004 n (%)	2005 n (%)	2006 n (%)	2007 n (%)	2007 n (%)
A1	117 (1.4)	157 (1.4)	8 (0.1)	5 (0.0)	2 (0.0)
A2	4 (0.0)	7 (0.1)	0 (0.0)	0 (0.0)	1 (0.0)
A3	1 (0.0)	6 (0.1)	6 (0.0)	2 (0.0)	0 (0.0)
B1	4 (0.0)	6 (0.1)	3 (0.0)	0 (0.0)	0 (0.0)
B2	1 (0.0)	0 (0.0)	1 (0.0)	0 (0.0)	0 (0.0)
C1	6 (0.1)	28 (0.2)	8 (0.1)	8 (0.1)	10 (0.1)
C2	908 (10.6)	1505 (13.3)	1241 (10.3)	981 (8.6)	1276(8.9)
C3	7 (0.1)	4 (0.0)	0 (0.0)	10 (0.1)	6 (0.0)
D1	991 (12.0)	914 (8.1)	722 (6.0)	801 (7.1)	444(3.1)
D2	119 (1.4)	154 (1.4)	222 (1.6)	215 (1.5)	20 (0.1)
D3	98 (7.7)	81 (5.4)	37 (2.4)	296 (36.2)	48(34.0)

Analysis of the table indicates high prevalence of D3 and D1 coded types of errors throughout the period under study ranging from 2.4% to 36.2% (cases missing date of specimen receipt in laboratory) and 3.1% - 12.0% (cases with missing ITD results) respectively. The next error category which had high incidence was group C2 (cases with inadequate interval between onset and date of stool specimen taken) with incidence rate of 8.6% - 13.3%. The results also show

errors against other codes: A1 code (wrongly formulated EPID number) - 0.1% - 1.4%; B1 (cases with missing date of onset) - 0.1% - 0.2%, D2 (cases with missing primary isolation results) - 0.1% – 1.6%. As for code B2 (mismatch of dates of onset and database), remarkably there were no errors identified over a five year period. Although errors in group A1 decreased after 2006, the analysis showed their higher values compared to other types of errors within group A.

4.5 Desegregation of most Prevalent Errors in the Clinical Database by Country and by IST

The following three tables (5-7) represent the number of cases with missing specimen receipt dates in the clinical database (error code: D1) by country in the sub-region for the five year period.

Table 5: Distribution of cases with missing specimen receipt dates in the countries of the IST West (Error code D1)

Country	2004	2005	2006	2007	2008	Total
Burkina Faso	58	51	43	18	99	269
Cote d'Ivoire	5	4	4	203	1	217
Guinea	79	48	6	37	42	212
Mali	61	79	7	3	0	150
Senegal	0	0	145	2	0	147
Níger	20	21	21	55	9	126
Sierra Leone	33	27	0	11	19	90
Liberia	15	50	10	10	0	85
Nigeria	0	0	1	50	13	64
Togo	3	36	9	0	12	62
Mauritania	9	16	9	12	2	48
Gambia	3	12	3	3	4	25
Algeria	16	0	0	0	7	24
Ghana	1	1	0	12	6	20
Guinea-Bissau	4	0	0	3	2	9
Cape Verde	1	0	3	1	1	6
Benin	1	0	0	1	0	2
Total IST West	309	347	261	421	217	1555

Burkina Faso, Cote d'Ivoire, Guinea had the highest number of cases 269 (17.2%), 217 (13.9%) and 212 (13.5%) respectively of the total number of errors (1555) in IST West. The first 5 countries in the table contributed 64% of all errors in IST West. Benin, Cape Verde and Guinea-Bissau had the least number of errors. The total number of D1 errors was fluctuating during the study period. Most of the errors- 421 (26.9%) were identified in 2007.

Table 6: Distribution of cases with missing specimen receipt dates in the countries of the IST Central (Error code D1)

Country	2004	2005	2006	2007	2008	Total
Chad	119	162	7	2	6	296
Central African Republic	120	126	0	0	0	246
Congo	35	53	3	2	0	93
Republic Democratic of Congo	0	5	30	34	0	69
Angola	11	0	15	32	1	59
Rwanda	6	7	21	6	7	47
Burundi	4	3	18	7	3	35
Gabon	7	8	0	0	1	16
Equatorial Guinea	6	4	0	2	0	12
Sao Tome and Principe	0	0	2	0	1	3
Total IST Central	308	368	96	85	19	876

Of the total number of errors recorded, 33% (296) were from Chad representing the highest number of cases with missing specimen receipt date in the clinical database. Chad and Central African Republic together contributed about 62% of all errors in the sub-Region. Sao Tome and Principe had the least number of errors (3). There was a remarkable decrease in D1 code errors throughout the study period from 308 cases in 2004 to just 19 cases in 2008.

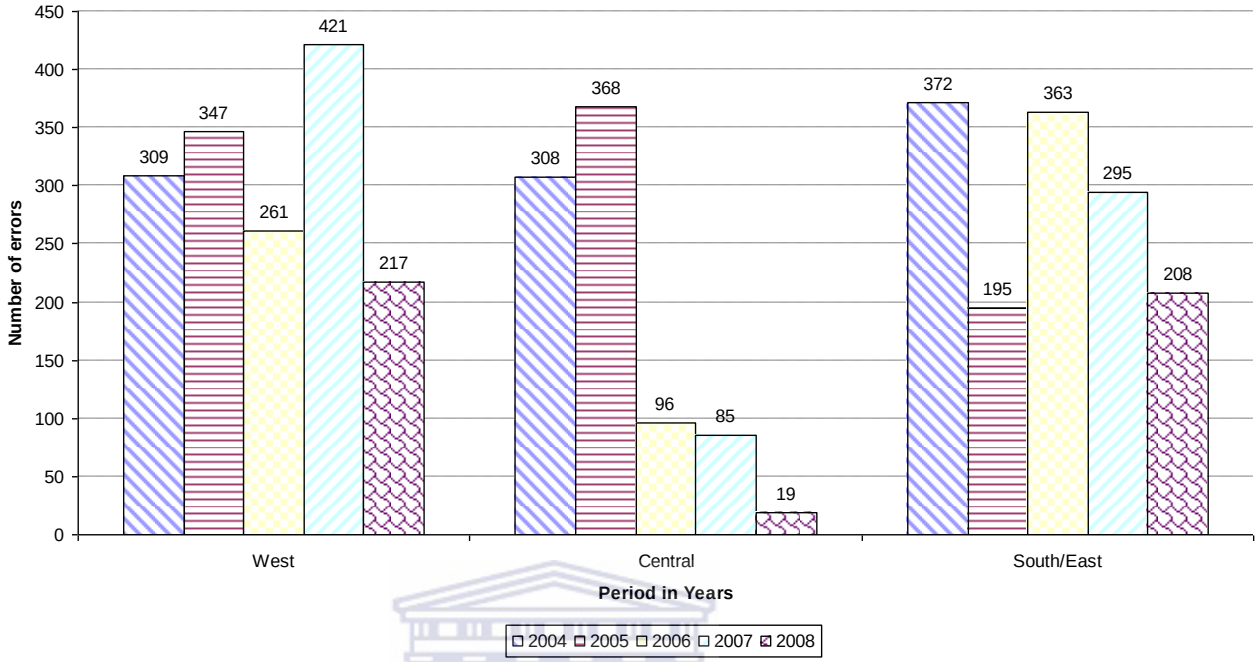
Table 7: Distribution of cases with missing specimen receipt dates in the countries of the IST South/East (Error code D1)

Country	2004	2005	2006	2007	2008	Total
Namibia	5	14	273	31	90	414
Tanzania	18	57	24	135	5	244
South Africa	193	14	2	5	2	216
Zimbabwe	95	19	7	34	18	174
Malawi	35	28	18	22	33	136
Eritrea	2	49	10	16	32	109
Mozambique	2	0	3	15	21	41
Lesotho	3	3	15	12	0	33
Kenya	16	5	2	3	2	28
Zambia	0	1	5	7	3	16
Botswana	0	4	0	5	0	9
Madagascar	3	1	0	4	1	9
Ethiopia	0	0	4	4	0	8
Comoros	0	0	0	2	1	3
Total IST South/East	372	195	363	295	208	1433

In IST South/East, Namibia had very high rate of D1 errors - 28.8% (414). The first five countries in the table had more than 80% of all errors recorded by countries in this IST. The last 5 countries in this IST had the lowest levels of D1 code errors ranging from 3 (Comoros) to 16 (Zambia).

As shown in Figure 5, there was in general a reduction in the prevalence of errors in the clinical database for all three ISTs from 2004 – 2008; reduction by 93.8% in IST Central, 29.8% - IST West and 44.1% in IST South/East.

Figure 5: D1 code errors by sub-region (ISTs) and year for the countries of the WHO African region



A closer look at the number of errors in Figure 5 over the period of time shows that:

IST West: The trend in number of errors showed a fluctuation of between 309 in 2004 and 347 in 2005, followed by a drop to 261 in 2006. This was followed by a sharp rise the following year to 421 and then reaching 217 in 2008. The countries with relatively high numbers of errors in 2004 were Guinea (79), Mali (61), Burkina Faso (58), Sierra Leone (33) and Niger (20). Nigeria, with the largest population and also the only endemic country of the region and with on-going polio virus transmission (CDC, 2011) over the period, significantly did not have any D1 code errors in 2004.

IST Central: There was an increase in errors from 308 in 2004 to 368 in 2005, followed by very significant and steady decline for the rest of the years to 96 and 85 (2006-07) and finally to 19 in 2008.

IST South East: There was a fluctuation similar to the data for IST West from 372, 363, and 208 in 2004, 2006 and 2008 respectively. This was due to the large numbers of errors in two countries, Namibia with a very small population with 273 in 2006 and Tanzania with 135 errors in 2007.

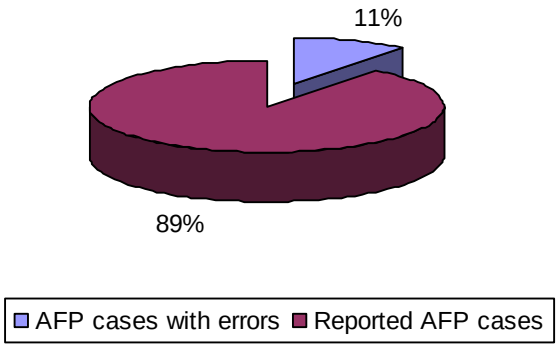
4.6 Errors in the Laboratory Database

The laboratory database consisted of 59,843 records of AFP cases for the period 2004 - 2008. These records represent data compiled from the 16 national and 10 WHO Reference laboratories for ITD results.

The number of errors in this database was 6,383 representing 11% of the total errors as shown in Figure 6.

Figure 6

AFP cases and total errors identified in the laboratory database



Errors per identified codes are presented in table 9 which shows frequency of various errors within major codes A, B, C, and D for the period under study. Compared with the clinical database, these proportions of errors in general are lower and in some error areas, such as B1, B2, and D1 no errors were recorded.

Table 8: Number and Proportion (%) of Errors in the Laboratory Database, African Region, 2004 – 2008*)

Type of error	2004 *n*(%)	2005 *n*(%)	2006 *n*(%)	2007 *n*(%)	2008 *n*(%)
AFP cases in	9761	11490	12152	12186	14254
Positive samples	1346	1405	1399	775	1300
A1	7 (0.1)	10 (0.1)	24 (0.2)	12 (0.1)	13 (0.1)
A2	41 (0.4)	37 (0.3)	231 (1.9)	34 (0.3)	38 (0.3)
A3	9 (0.1)	6 (0.1)	41 (0.3)	14 (0.1)	12 (0.1)
B1	0	0	0	0	0
B2	0	0	0	0	0
C1	17 (0.2)	22 (0.2)	22 (0.2)	17 (0.1)	4 (0.0)
C2	899 (9.2)	1381(12.0)	1103(9.0)	933 (7.7)	1078 (7.6)
C3	N/A	N/A	N/A	N/A	N/A
D1	0	0	0	0	0
D2	1 (0.0)	5 (0.0)	2 (0.0)	42 (0.3)	233(1.6) D3
25 (1.9)	23 (1.6)	31 (2.2)	0	12 (0.9)	

*) The coding of error type is the same as the one used for the clinical database.

The most prevalent errors were: A2-incomplete EPID numbers, A3-discrepant year in EPID number and date of onset, C1-cases with negative interval between date of onset and the date stool sample was collected, D3 - isolates missing ITD results and C2 - cases with interval between onset of paralysis and date 2nd stool was collected. Error code C2 had the highest incidence rate throughout the five year period ranging from 7.6% (2008) to 12.0% (2005).

4.7. Distribution of Major Errors in the Laboratory Database

The following results show the distribution of the errors identified in the laboratory database from 2004 to 2008.

Table 9: Distribution of Cases with Incomplete EPID number (A2 error) by Laboratory, 2004-2008

Country	2004	2005	2006	2007	2008
	n(%)	*n*(%)	*n*(%)	*n*(%)	
Central African Republic	2 (4.9)	6 (16.2)	0	0	0
Cote d'Ivoire	0	0	0	1 (0.4)	0
Kenya	1(2.4)	0	0	2 (5.9)	4 (10.5)
Nigeria (Maiduguri)	0	0	0	0	1 (2.6)
Democratic Republic of Congo	0	2 (5.4)	2 (0.9)	0	0
Senegal	0	0	0	0	1 (2.6)
South Africa	36 (87.8)	24(64.9)	217 (93.9)	31(91.2)	30 (78.9)
Uganda	1 (2.4)	1 (2.7)	4 (1.7)	0	0
Zambia	0	2 (5.4)	0	0	0
Zimbabwe	1 (2.4)	2 (5.4)	7 (3.0)	1 (2.9)	2 (5.3)
Total	41	37	231	34	38

Laboratories in Algeria, Cameroon, Ethiopia, Ghana, Nigeria, and Madagascar did not have A2 type error for the five year period. The highest number of total cases with A2 error was recorded in 2006 (231) while the lowest number - in 2007 (34 cases). The laboratories that consistently had this error over a five year period were South Africa and Zimbabwe with former having highest proportion of them each year. Cote d'Ivoire had a single error in 2006 during the entire period under review.

Table 10: Distribution of cases with discrepant year in EPID number and date of onset of paralysis (A3 error code) by laboratory, 2004-2008

Country	2004	2005	2006	2007	2008
	n(%)	*n*(%)	*n*(%)	*n*(%)	
Cameroon	0	0	2(4.9)	0	0
Cote d'Ivoire	0	0	1(2.4)	0	1(8.3)
Ethiopia	0	0	0	1(7.1)	0
Ghana	1 (11.1)	0	7(17.1)	4(28.6)	2(16.7)
Kenya	1 (11.1)	0	1(2.4)	3(21.4)	0
Madagascar	0	1(16.7)	0	0	0
Nigeria (Maiduguri)	0	0	2(4.9)	1(7.1)	1 (8.3)
Democratic Republic of Congo	0	0	2(4.9)	0	1 (8.3)
Senegal	0	0	8(19.5)	0	0
South Africa	1 (11.1)	1(16.7)	10(24.4)	4(28.6)	0
Uganda	1 (11.1)	1(16.7)	4(9.8)	0	0
Zambia	3 (33.3)	2(33.3)	4(9.8)	0	0
Zimbabwe	2 (22.2)	1(16.7)	0	1(7.1)	2 (16.7)
Total	9	6	41	14	7

Laboratories from Algeria and Central African Republic did not have A3 error. The highest number of cases (41) with such discrepancies was recorded in 2006 while the least (6) was recorded in 2005. The laboratories that had errors for most of the years were South Africa, Ghana, and Zimbabwe. South Africa had highest proportion of discrepancies each year ranging from 11.1% in 2004 to 28.6% in 2007. Cameroon had the least number of errors (4.9%) in 2006.

Table 11: Distribution of cases with negative interval between date of onset and date stool specimen collection (C1 error code) by laboratory, 2004 – 2008.

Country	2004	2005	2006	2007	2008	Total
	n(%)	*n*(%)	*n*(%)	*n*(%)		
Cameroun	1(5.9)	2(4.5)	0	0	0	2
Cote d'Ivoire	1(5.9)	0	0	0	0	1
Ghana	0	2(9.1)	0	0	0	2
Nigeria (Ibadan)	2(11.8)	0	12(54.5)	7(31.8)	0	21
Nigeria (Maiduguri)	3(17.6)	2(9.1)	0	0	0	5
Kenya	1 (5.9)	2(9.1)	1(2.4)	0	0	4
Democratic Republic of Congo	1(5.9)	0	0	4(23.5)	0	5
Senegal	4(23.5)	0	1(4.5)	3(17.6)	0	8
South Africa	0	0	12(54.5)	3(17.6)	1(25.0)	17
Uganda	1(5.9)	0	0	1(5.9)	0	2
Zambia	3(17.6)	2(9.1)	1(4.5)	5(29.4)	3(75.0)	17
Zimbabwe	0	1(4.5)	0	1(5.9)	0	2
Total	17	22	22	17	4	86

Laboratories from Algeria, Ethiopia, and Madagascar did not have C1 code error. Most of these errors were recorded in 2005 and 2006 with both years having 22 errors while in 2008 the number of errors dropped to 4. The highest number of errors over a five year period (21) was by one laboratory in Ibadan, Nigeria while Cote d'Ivoire had the least (1). The laboratory in Zambia had C1 code errors throughout the 5-year period.

The following table represents errors in one of the important areas of polio surveillance: date of onset of AFP and date on which stool sample was collected which should be within 14 days following the onset of paralysis (C2 code).

Table 12: Distribution of cases where date of onset and date stool sample collected was greater than 14 days by laboratory, 2004-2008 (Error code C2) for all the laboratories in the listed countries

Country	2004	2005	2006	2007	2008
	n(%)	*n*(%)	*n*(%)	*n*(%)	
Reported AFP cases	9761	11490	12299	12186	14254
Algeria	8 (0.9)	16 (1.2)	13 (1.2)	28 (3.0)	35 (3.2)
Cameroon	61(6.8)	55(4.0)	29(2.6)	53(5.7)	64(5.9)
Central African Republic	18(2.0)	11(0.8)	6(0.5)	5 (0.5)	6 (0.6)
Cote d'Ivoire	70(7.8)	124 (9.0)	30(2.7)	39 (4.2)	51 (4.7)
Ethiopia	68(7.6)	167(12.1)	65(5.9)	98(10.5)	119(11.0)
Ghana	24(2.7)	16 (1.2)	15(1.4)	10(1.1)	29(2.7)
Nigeria (Ibadan)	211(23.5)	282(20.4)	164(14.9)	79(8.5)	127(11.8)
Nigeria (Maiduguri)	188(20.9)	315(22.8)	339(30.7)	120(12.9)	180(16.7)
Kenya	51(5.7)	29(2.1)	58(5.3)	59(6.3)	50 (4.6)
Madagascar	13 (1.4)	25(1.8)	15 (1.4)	10 (1.1)	180(16.7)
Republic Democratic of Congo	47(5.2)	153(11.1)	176(16.0)	256(27.4)	195(18.1)
Senegal	54(6.0)	66(4.8)	20(1.8)	8(3.0)	52(4.8)
South Africa	30 (3.3)	46(3.3)	87(7.9)	69(7.4)	74(6.9)
Uganda	22 (2.4)	42(3.0)	36(3.3)	54(5.8)	56(5.2)
Zambia	17 (1.9)	15(1.1)	24(2.2)	9(1.0)	19(1.8)
Zimbabwe	17(1.9)	19(1.4)	26(2.4)	16(1.7)	17(1.6)
Total	899(9.2)	1381(12.0)	1103(9.0)	933(7.7)	1078(7.6)

The highest number of reported cases with interval exceeding 14 days between date of onset and stool collection date was in 2005 (1381). Among the laboratories, the highest number of errors over a five year period was Maiduguri in Nigeria with 1142 errors while the laboratory in Central African Republic had only 46 errors.

4.8 Comparison of Extent of Different Errors in the Clinical and Laboratory databases

Further analysis refers to discrepancies between clinical and laboratory databases illustrated in the series of graphs below.

Figure 7

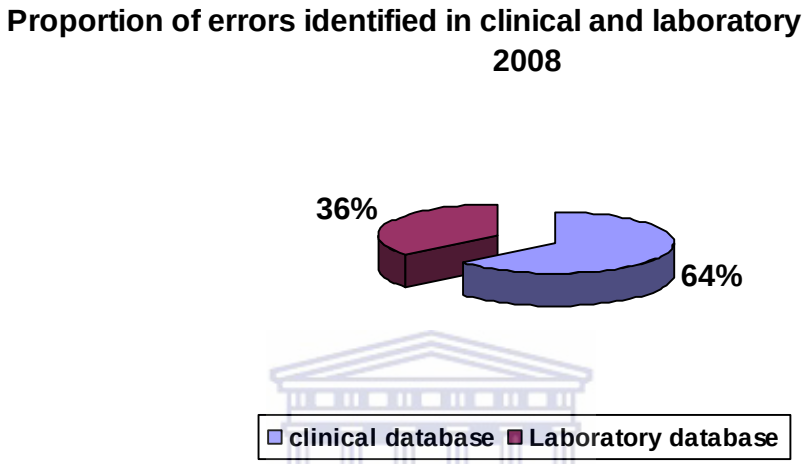
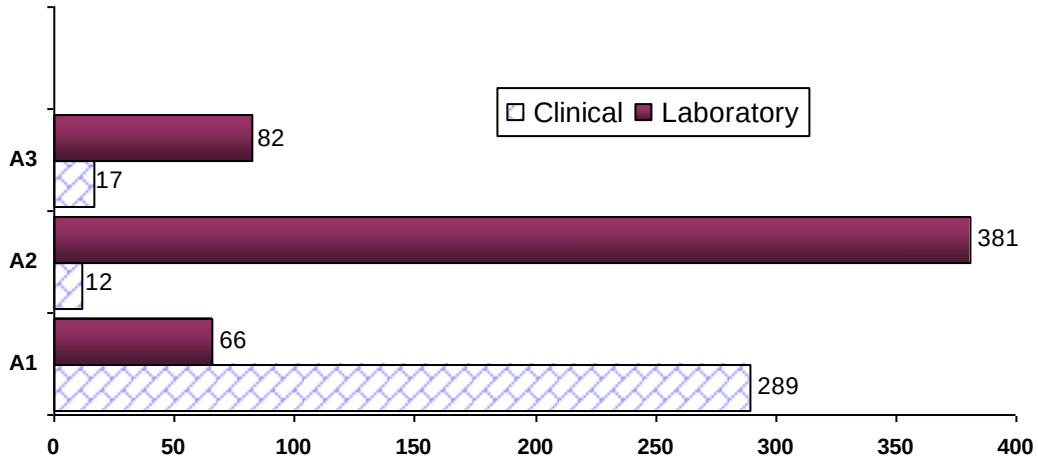


Figure 7 indicates that the number of errors in the clinical database far exceeded those in the laboratory database representing 64% of the total.

To find out what the differences in recorded error types between clinical and laboratory databases were the errors were analysed per major groups as shown in the following graphs.

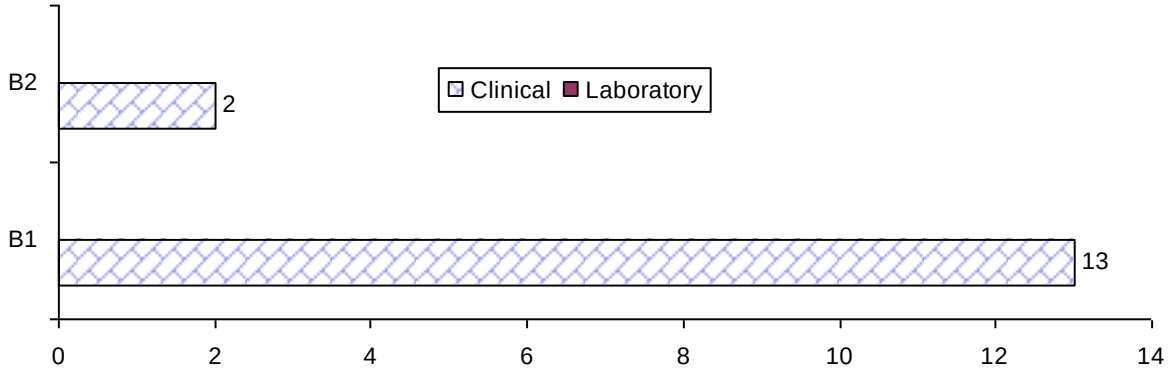
Figure 8: Number of AFP cases with the discrepancy with EPID number wrong formulation (A1), incomplete (A2) and mismatch of dates (A3).



The graph illustrates comparatively higher number of errors (four-fold) in the clinical database against A1 error code. As for A2 code referring to incomplete EPID number and A3 (mismatch of year in EPID numbers) most of the errors in both cases were identified in the laboratory database.

The next figure relates to error code block B on missing date of onset (B1) and mismatch of year in the databases (B2) showing that all errors were attributed to the clinical database with no errors recorded in the laboratory database.

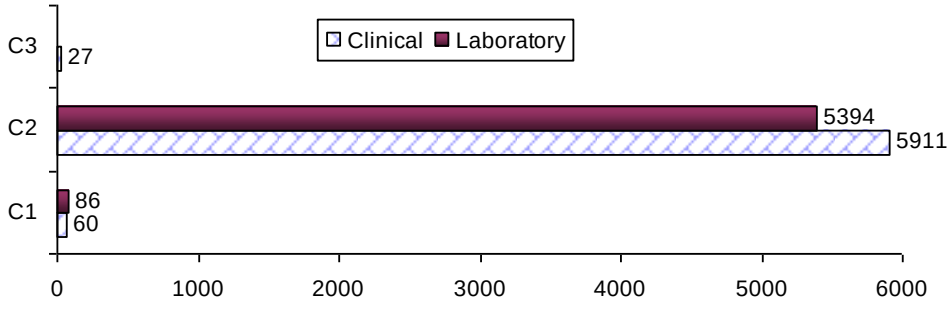
Figure 9: Errors in missing date of onset (B1) and mismatch of year of onset in the databases (B2).



There were 13 errors recorded against B1 error code and only 2 against B2 code.

Figure 10 below refers to error code block "C" on collection of stool samples from AFP cases - one of the important aspects of polio surveillance.

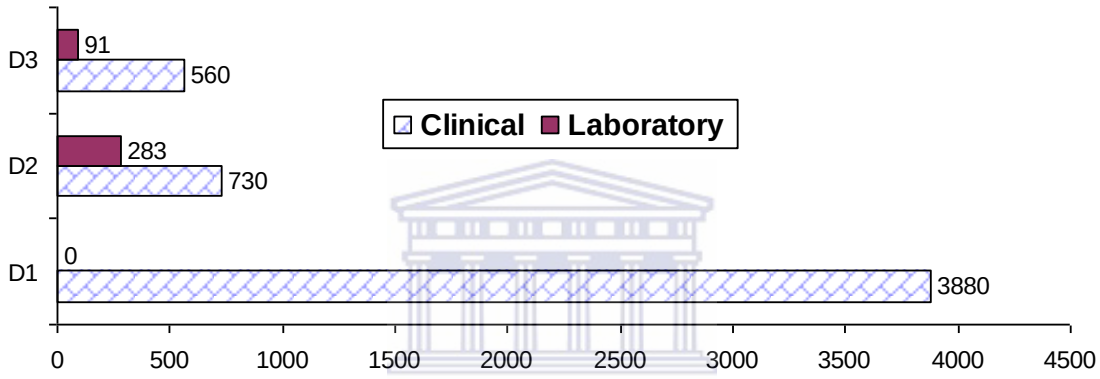
Figure 10: Number of AFP cases with discrepancies in date of onset (C1), date stool sample collected (C2) and time between 2 stool samples (C3).



Analysis of the above figure shows comparatively low number of errors in C1 and C3 while C2 constituted the major error in both databases with a slight domination in the clinical database.

As shown in figure 11 D1errors were identified only in the clinical database with high frequencies (3880 error cases). Some errors in D2 and D3 were shared between the clinical and laboratory databases with increased numbers in the former.

Figure 11: Number of errors in block D code on cases missing dates specimen received in the laboratory (D1), missing primary isolation results (D2) and ITD results (D3)



4.9 Non-coded errors in databases

This analysis included discrepancies between inter-related databases as well as the timeliness of specimen delivery from the field to central and laboratory levels. These errors were not included in the coded system categorized in table 4.

4.9.1 Discrepancies in reported AFP cases and laboratory results between databases

The table below illustrates differences in total numbers of reported AFP cases and positive isolates in both databases.

Table 13: Reported AFP cases in the clinical and laboratory databases over the period 2004 to 2008

Type of error	2004 n	2005 n	2006 n	2007 n	2008 n
Clinical database					
Reported AFP cases	8560	11341	12067	11355	14296
AFP with stools	8537	11312	12034	11333	14263
Positive isolates	1270	1505	1562	721	1387
Laboratory database					
AFP cases in	9761	11490	12152	12186	14254
Positive isolates	1346	1405	1399	885	1300

The table above clearly indicates significant inconsistencies both in numbers of reported AFP cases which are more in the laboratory database, and number of positive isolates. Throughout the five year period, with the exceptions of 2004 and 2007, the data indicated that the number of positive isolates in the clinical database was more than those in the laboratory.

4.9.2 Timeliness of specimen movement from the field to the central level and from central level to the laboratory

The third objective of the study focused on determining the timeliness of reporting, the proportion of cases in the clinical and laboratory databases received within 7 days at the central level and in the laboratory. The timeliness in the clinical database was calculated using date the stool sample was collected and the date it was received at central level while the laboratory timeliness was calculated using the date when the sample was sent to the laboratory and the date when it was received in the laboratory.

Findings in the study showed that while the number of reported AFP cases steadily improved in the clinical database from 8560 in 2004 to 14296 in 2008, the timeliness has constantly been 50% or below compared to over 80% in the laboratory database (Table 12 and Figure 2).

Table 14: Proportion of specimens from AFP cases received within 7 days at central level and by the laboratory

Year	Clinical database		Laboratory database	
	Reported AFP cases	AFP received within 7 days	AFP cases in Laboratory	AFP received within 7 days
2004	8560	2346 (27.4%)	9761	8545 (88.5%)
2005	11341	4662 (41.1%)	11490	9794 (85.2%)
2006	12067	4561 (38.8%)	12299	10884 (88.5%)
2007	11355	4344 (38.3%)	12186	10209 (84.8%)
2008	14296	7156 (50.1%)	14254	12256 (86.0%)

Table 15: Reported AFP cases in the clinical and laboratory databases

Type of error	2004 n	2005 n	2006 n	2007 n	2008 n
Clinical database					
Reported AFP cases	8560	11341	12067	11355	14296
AFP with stools	8537	11312	12034	11333	14263
Positive isolates	1270	1505	1562	721	1387
Laboratory database					
AFP cases in	9761	11490	12152	12186	14254
Positive samples	1346	1405	1399	885	1300

Figure 12: Percentage of timely stool specimens' reception in laboratory (within seven days of shipment)

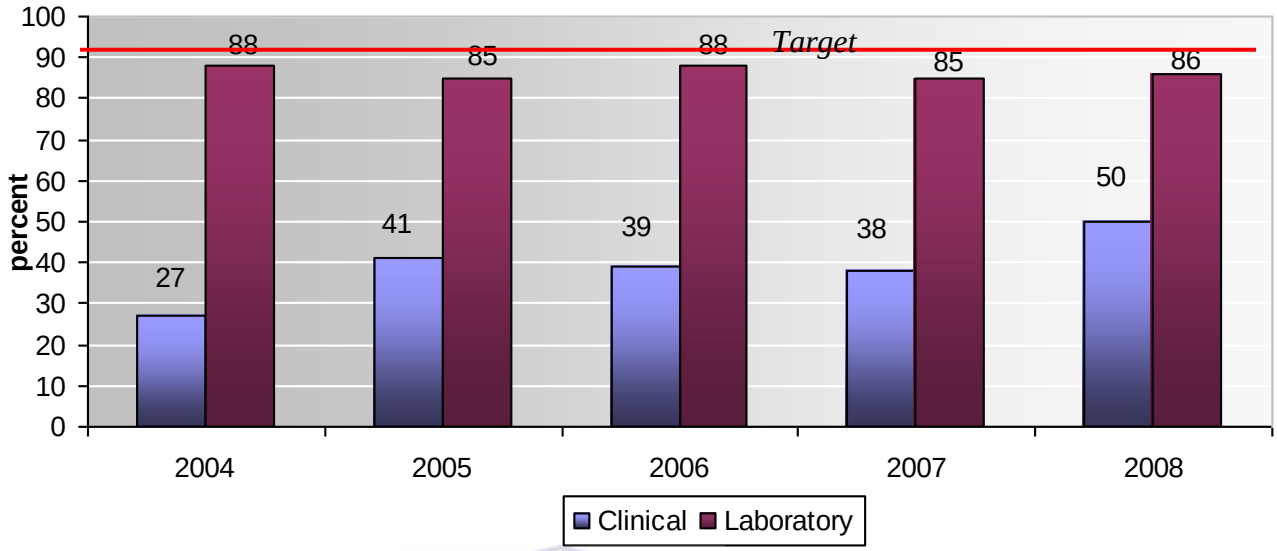


Figure 12 above shows that the numbers of specimen which were received in the laboratory and in time has almost remained unchanged, but well above the target of 80%, while timeliness of receipt of clinical data has increased significantly, almost two fold between 2004 and 2008, but without attaining the target of 80%. The increase from 27 in 2004 to 41 in 2005 was statistically significant ($P<0.005$)

CHAPTER FIVE

DISCUSSIONS

Clinical and Laboratory databases: Most of the study objectives were related to identification and quantification of errors in the clinical and laboratory databases maintained by the WHO, African Region. The current methods used for sharing data between the national AFP Surveillance programs and the inter-country laboratories may be plagued by poor data handling and transmission (de Gourville *et al.* 2006). Attempt to find a concise framework for identification of errors in the literature as a guidance to this study was not successful as there were few sources, especially related to errors in polio databases, in the African Region. Unlike reported vaccination coverage data which can be validated through independent surveys (Murray *et al*. 2003) the quality of AFP data can only be assessed through reviews of records or datasets such as the present study.

During this study an effort was made to develop an error framework with coding which included four major codes (A, B, C, D) and 11 sub codes, reflecting different elements of AFP surveillance. The error framework was informed by the knowledge of the variety of errors commonly encountered in AFP databases over the past years and also contributes significantly to the surveillance indicators. There is no doubt that this framework implying precise error categories if adapted throughout the African Region could facilitate a uniform approach in identifying data gaps and advising countries to rectify common errors thus improving overall AFP surveillance. In addition, the author also reviewed non-coded errors referring to inconsistencies of AFP cases reported and timeliness of stool specimens' delivery from the field to central and laboratory levels.

The study reviewed a total of 57,619 clinical and 59,843 laboratory AFP cases reported during the period 2004 – 2008 by 46 member states of the WHO Regional Office for Africa. The study assessed the common errors in both databases and their implications on polio eradication programmes in the WHO African Region.

Of the eleven types of errors identified, those errors relating to cases with pending laboratory results and cases with intervals between date of onset and stool collection or date specimen was received in the laboratory were found to be the most prevalent in both databases. The proportions of these two common types of errors were higher in the clinical database ranging from 7.7% in 2004 to 58.9% in 2008 for cases missing intra-typic differentiation results compared to proportions in the laboratory database that had a range of 0.0% to 1.9% over the same period. Cases where the date of onset and date of stool collection was greater than 14 days was higher in both databases with slightly higher rates in the clinical database ranging from 8.6% in 2004 to 13.3% in 2008 compared to 7.6% and 12.0 in the laboratory database. These results confirm the earlier findings of inconsistency of immunization data at point of capture and after transfer of the data (Mavimbe, Braa and Bjune, 2005). Such errors as revealed by the present study call for action to be taken at the point of data capture or when data is transferred from one point to another, to minimize the introduction of errors.

Critical indicators: Regarding the surveillance standard on annualized AFP rate (standard: 1 AFP case per 100,000 children < 15 years of age), the present study has revealed that the African Region exceeds the target for this key indicator by several times, reporting 2.9 – 4.5 per 100,000 children <15 years AFP cases throughout 2004-2008. Moreover, the Region also did well as for other important surveillance indicators requiring the collection of two stool specimens per an AFP case: 99% against the target of 80%. Based on these findings, the hypotheses 1 (section 1.6)

assuming that the AFP surveillance in the African Region show a sub-standard performance is rejected.

Auditing AFP Surveillance Errors

(a) **Coded errors in clinical database**: The analysis by major error groups in clinical database showed that in groups A (errors in EPID numbers) and especially group B (errors in date of onset) are not high (0.1-1.4%) compared with the other two groups (C and D). In group C, the dominant error group was C2 (stool sample taken beyond 14 days of onset) contributing 13.3% of all errors in 2005, reduced to 8.6% in 2008 which should also be considered as high. The importance of this error is enormous as the likelihood of isolating virus in the stool is greatest if the specimens are collected within 14 days of onset of paralysis and that it diminishes markedly by 30 days (WHO, 2001). This finding may explain why some WPV cases are not detected in time to initiate supplementary immunization until there has been a spread to other parts of the country or to other countries, resulting in the presence of orphan viruses and constituting a setback to the programme in the African region.

The group D errors, having greater proportion among all errors categories, are indicative of poor interaction between those involved in compiling the clinical and laboratory databases. These errors refer to missing dates the laboratory received the samples (31.1%) as well as primary or ITD results submitted by laboratories to clinical databases (0.1-1.5%) and 2.4-34.,0%) respectively during the various years of the study. The sharp increase of D3 errors on ITD results in 2007 and 2008 may have been caused by changes in laboratory algorithm in 2007 reducing deadlines of submission of laboratory results from 28 to 14 days for primary isolation and after

primary results from 21 to 7 days for ITD results (AFRO memorandum to laboratories and countries, 2007).

The analysis of the D1 error type by WHO sub-region and country indicates that countries in IST Central have the least number of errors with a significant reduction during 2004-2008 (from 308 errors in 2004 to 19 in 2008) while the performance of the countries within the other two ISTs needs substantial overhaul due to high prevalence of D1 errors, not showing a meaningful decline during the study period. The most three D1 "error-infested" countries in IST West were Burkina Faso, Cote d'Ivoire, Guinea; in IST Central: Chad, Central African Republic, Democratic Republic of Congo and in the IST South/East: Namibia, Tanzania and South Africa. Based on the above discussion, the hypothesis 2 regarding inadequacy of AFP surveillance indicators in clinical database, is confirmed. One may speculate that the differences between countries and ISTs in terms of numbers of errors may be attributable to the differences in qualifications and the levels of experience of the staff in different countries as well as the workload. However since the present study did not specifically determine staff numbers per facility, qualifications and level of experience, it is impossible to assign specific reasons for the observed differences. Fundamentally errors in data can be attributed to the inability to accurately record, report, transmit and manage data (Bosch-Capblanch *et al.* 2009). At any point along the chain errors can be introduced, either through omissions, wrong entries, delays in transmission or combinations of these. Other studies have shown that the qualification of staff as well as the workload had very significant, direct influence on the quality of data collected at health facilities (Makombe *et al.* 20008). It is plausible that availability of staff, qualifications and workload may have contributed to the errors observed in the present study.

(b) **Coded Errors in Laboratory databases**: In this database there are few errors related to wrong or incomplete formulation of EPID numbers and mismatch of dates (error code A) between EPID number and the date of onset. However, errors in C2 code were high amounting to 7.7% - 12.0% during the study period. These errors are related to the key surveillance indicators such as the date of onset of paralysis and stool sample collected within 14 days of onset of paralysis. As discussed above, errors in this category can compromise the detection of the wild poliovirus in a sample. Relatively, more errors were recorded against error code D3 related to missed ITD results (0.9-2.2%) which may be related to the reluctance of staff to ensure the timely entry of results into the database or shortage of staff dealing with data. Indeed, in most laboratories, the technician who carries out the tedious laboratory analysis is also often responsible for entry of the results into the database. Engagement of data entry clerks with sole responsibility for this task could minimize the number of errors. Earlier studies (Harris *et al.* 2003, Danovaro *et al.* 2007) referred to errors in data attributable to inadequate performance by demotivated personnel and high staff workload. Comparison of the frequency of the most prevalent errors by laboratories revealed that most errors related to code C2 were dominant, contributing up to 12.0% of all identified errors. Two laboratories located in Ibadan and Maiduguri in Nigeria, contributed approximately from 22% to 45% of errors among laboratories in the African Region. Based on this data, hypotheses 3, on lower rates of errors in the laboratory database as not being significant, is rejected. By paying attention to these two laboratories and carrying out in-depth reviews the causes of the errors could be identified and quick remedial action taken to improve data quality and ensure polio eradication.

c) Non-coded errors in clinical and laboratory databases

Inconsistencies of AFP records: Maintaining high data quality requires going through the data periodically, updating, standardizing and eliminating duplicates to create a single view of the data, even if it is stored in multiple disparate systems (WHO, 2001).

The table 15 in results section clearly indicates significant inconsistencies both in numbers of reported AFP cases which are more in the laboratory database, and number of positive isolates. The data, throughout the five year period, with the exception of 2004 and 2007, indicated that the number of positive isolates in the clinical database was more than those in the laboratory database.

The reason for mismatch of > 2300 AFP cases between the clinical and laboratory databases is not well understood although it could be attributed to some specimens being sent directly to the laboratory, thus bypassing the clinical database at national level. In a well-functioning surveillance system there should be cross-communication between two databases or those who manage them for harmonization of data. This aspect is not sufficiently addressed by existing reports to locate specific sources of this mismatch, and could be considered for further research at country level. It is also likely that the required periodic updating, standardizing and eliminating duplicates as recommended (WHO, 2001) is not being carried out. Elimination of such disparities is critical in ensuring that the goals of the GPEI are achieved and the region is certified free of poliomyelitis.

Timeliness: Timeliness of shipment of specimen from field to laboratory and between laboratories is a significant factor in determining the outcome of the final laboratory results. Findings in the present study showed that while the number of reported AFP cases steadily improved in the clinical database from 8560 in 2004 to 14296 in 2008, the timeliness constantly remained at 50% or below compared to over 80% for the laboratory database. This poor timeliness is an indication of possible problems in transportation of specimens from the field to the central level (WHO/Geneva, 2001). The implication of this delay in specimen transportation is that new WPV cases are not readily identified and in time to initiate the required supplementary immunization activities in line with the PEI strategy so as to eradicate the disease in Africa.

d) Implications of identified errors on GPEI

The present study identified the existence of errors against all of 11 error codes in both the clinical and laboratory databases, which is widespread in countries of the three ISTs of the WHO African region. In general, all these errors have implications for AFP surveillance and consequently the Global Polio Eradication Initiative, since they compromise the accurate identification of infection with new WPV, information which is critical for the planning and implementation of vaccination campaigns in order to prevent the spread of the virus, ultimately the disease and attainment of certification of African Region as polio free as has been done in the South American and European Regions (de Quadros, 1997).

Some of these implications are described in the table below related to each of the errors identified during the study, and their importance is indicated as per assessment scale.

While the present study assessed the quality of data and its implications to the polio eradication initiative in the African region, the study could not establish or assign accurate reasons for the errors which were identified since the study was retrospective. It was not possible to establish the timeline for correction of errors. Additional issues such as staff qualifications, experience and workload could not be ascertained and could be the subject of future studies.

CHAPTER SIX

CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions

The study reviewed a total of 57, 619 clinical and 59, 843 laboratory AFP cases reported during the period 2004 – 2008 by 46 Member States of the WHO Regional Office for Africa to find the extent of errors in both databases and their implications on polio eradication programmes in the WHO African Region.

The study revealed that the African Region exceeds the target of 1 AFP/100,000 children *<15* years indicator required by the GPEI by several times reporting 2.9–4.5 per 100,000 children <15 years AFP cases throughout 2004-2008. Moreover, the Region did also well as for other important surveillance indicators requiring the collection of two stool specimens per an AFP case: 99% against the target of 80%.

The analysis by major error groups in clinical database showed that the proportion of records with errors in groups A and B are not high (0.1-1.4%) compared with other two groups (C and D). In group C, the dominant error group was C2 related to optimal time of specimen collection contributing 13.3% of all errors in 2005, reduced to 8.6% in 2008.

The group D errors referring to missing date of specimen receipt in laboratory, having greater proportion among all errors categories, are indicative of poor interaction between clinical and laboratory databases or managers of these databases.

The analysis of the D error type related to laboratory results by WHO sub-region and country indicates that countries in IST Central have least number of errors with a significant reduction from 308 to 19 errors while the performance of other two ISTs needs substantial overhaul due to high prevalence of D1 errors, not showing a meaningful decline during the study period.

Comparison of the frequency of most prevalent errors by laboratories, revealed that most errors are related code C2 were dominate contributing up to 12.0% of all identified errors. Two laboratories located in Ibadan and Maiduguri in Nigeria, contributed approximately from 22% to 45% of errors among laboratories in the African Region. Based on this data, hypotheses 3, on lower rates of errors in the laboratory database as being not significant, is rejected.

Findings in the study showed that while the number of reported AFP cases steadily improved in the clinical database from 8560 in 2004 to 14296 in 2008, the timeliness of entering lab results has constantly been 50% or below compared to over 80% in the laboratory database.

The study identified existence of errors against all of 11 error codes in both the clinical and laboratory databases. There are more errors in the clinical database than in the laboratory database. These errors have implications for the Global Polio Eradication Initiative with some of them being critical for certification of African Region as polio free.

6.2 Recommendations

The following recommendations are proposed based on this study result:

- Countries should improve on the capture of clinical data by engaging qualified staff, who should be well trained and motivated.
- Countries are urged to improve mechanisms to clean and harmonize data between clinical and laboratory databases with a view to improving the quality of data. This is one of the conditions for attracting more support towards polio free certification status.
- Targeted training through workshops or traveling seminars for staff working with data in both clinical and laboratory databases are recommended.
- Organize cluster meetings or focus group discussions with data managers in countries with persistent problems or errors in their databases.
- Organize exchange visits among data managers to share best practices in data handling.
- Make presentations on the findings of this study during the annual EPI managers' meetings in all three ISTs
- What about the meeting of managers of the clinical and laboratory databases on frequent bases.
- Where dedicated data entry staff are unavailable as is the case in most laboratories, well trained and motivated data entry staff should be engaged to enter laboratory data.

REFERENCES

AbouZahr C. and Boerma T. (2005). Health information systems: the foundation of public health. *Bulletin of the World Health Organization*, 83 (8).

Basu R.N. (1982). Expanded programme on immunization and primary health care. Journal of communicable Disease, September; 14(3):183-188.

Birmingham M.E., Linkins R.W., Hull B.P. and Hull H.F (1997). Poliomyelitis Surveillance: The Compass for Eradication. The Journal of Infectious Diseases. 175(Sippl 1):S146-50.

Biritwum RB, Devres JP, Ofosu-Amaah S, Marfo C, Essah ER. Prevalence of children with disabilities in central region, Ghana. West African Journal of Medicine; 20(3):249-55.

Bosch-Capblanch, X, Ronveaux O, Doyle V, remedies V, Bchir Abdallah. (2009). Accuracy and quality of immunization information systems in forty-one low income countries. Tropical Medicine and International Health 14; 2-10

de Quadros, CA (1997). Global Eradication of Poliomyelitis and Measles: Another Quiet Revolution. Annals of Internal Medicine 12792):156 -158

Centers for Disease Control and Prevention (2006). *Advanced Management and Analysis of Data Using Epi Info for Windows*. Division of Global Public Health Capacity Development, Centers for Disease Control and Prevention, Atlanta Georgia, USA. Available from: http://www.cdc.gov/cogh/dgphcd/training/pdf/avanced_Management_manual.pdf.

Centers for Disease Control and Prevention (2010). Progress toward interruption of wild poliovirus transmission – worldwide 2009. Morbidity and Mortality Weekly Report 59(18) 545-550

Centers for Disease Control and Prevention (2011). Progress toward poliomyelitis eradication --- Nigeria, January 2010--June 2011. Morbidity Mortality Weekly Report 60:1053-7.

Danovaro C., Gilani Z., and Burton A. (2007). *The Quality of Coverage Data: A Regional Overview.* Washington D.C: Pan American Health Organization (PAHO).

de Kedt E. (1989). Making health policy management intersectoral: Issues of information analysis and use in less developed countries. *Social Science and Medicine,* 4: 503-514.

de Gourville E., Duintjer T. RJ., Sangrujee N., et al. (2006). Global surveillance and the value of information: the case of the global polio laboratory network. *Risk Analysis,* 26(6): 1557-69.

De Gouville E. and Dowdle WR. (2004). The role of surveillance in polio eradication and identification of emerging viral encephalitis. Weekly epidemiological record. No.18 13-20.

Global Polio Eradication Initiative (GPEI) (2008). *Poliomyelitis – The disease and the virus.* Accessed online: www.polioeradication.org/disease.asp [16/12/08].

Harris BN, Durrheim DN, Ogunbanjo GA (2003). Polio eradication—the validity of surveillance indicators. *Tropical Medicine International Health,* 8:386–391.

Health Protection Surveillance Centre (2007). *Acute Flaccid Paralysis Surveillance*. Dublin: Health Protection Surveillance Centre. Available from: http://www.hpsc.ie/hpsc/A-Z/VaccinePreventable/Polio/Guidance/File,2461,en.pdf.

Jani JV, Jani IV, Araujo C, Sahay S, Barreto J and Bjune G (2006). Assessment of routine surveillance data as a tool to investigate measles outbreak in Mozambique. *BMC Infectious Diseases,* 6: 29. Available from: http://www.biomedcentral.com/147-2334/6/29

Kew OM, Sutter RW, de Gourville EM, Dowdle WR, Pallansch MA (2005). Vaccine-derived polioviruses and the endgame strategy for global polio eradication. *Annual Review of Microbiology*, 59: 587-635.

Lim SS, Stein DB, Charrow A, Murray CJ (2008). Tracking progress towards universal childhood immunisation and the impact of global initiatives: a systematic analysis of three-dose diphtheria, tetanus, and pertussis immunisation coverage Dec 13;372 (9655):2031-46

Macfarlane S. (2005). Harmonizing Health Information Systems with Information Systems in other Social and Economic Sectors. *Bulletin of the World Health Organization*, 83 (8): 561-640.

Makombe S, Hochgesang M, Jahn A, Tweya H, Hedt B, Chuka S, Kwong-Leung YJ, Aberle-Grasse J, Pasulani O, Bailey C, Kamoto K, Schouten EJ, and Harris AD (2008). Assessing the quality of data aggregated by antiretroviral treatment (ART) clinics in Malawi. *Bulletin of the World Health Organization*, 86(4): 310-314

Mavimbe J, Braa J, and Bjune G (2005). Assessing Immunization Data quality from routine reports in Mozambique. *BMC Public Health*, 5:108.

Murray CJL, Bakhuti S, Gupta N, Saba M, Tandon A, Thieren M. (2003). Validity of reported vaccination coverage in 45 countries. The Lancet 362: 122-1026.

Murray C., J.L, Lopez A.D, Mathers C.D., 2004. *The Global Epidemiology of Infectious Diseases*. :(2004) Ed. Murray C. J.L., Lopez A.D., and Mathers C.D. Chapter 5, Poliomyelitis by Shibuya K. Murray C.J.L.; Page 111-149

Nokleby H. et el, (2010). European Surveillance Disease Bulletin. 29 April 2010;15(17). Pii:19556.

Nsubuga P., McDonnel S., Perkins B., Sutter R., Quick L., White M., Cochi S., Otten M. (2002). Polio eradication initiative in Africa: influence on other infectious disease surveillance development. BMC Public Health; 2:27-35.

Pavlov DN et al (2005). *Prevalence of vaccine-derived polioviruses in sewage and river water in South Africa*. Water Research Volume 14: 3309-3319

Polionews No.32, 1st Quarter, 2009 EDITION. WHO/CDC/UN IEF (2009)

Ronveaux O, Rickert D, Hadler S, Groom H, Lloyd J, Bchir A, Birmingham M. 2004 The immunization data quality audit: verifying the quality and consistency of immunization monitoring systems. Bulletin World Health Organization 83: 503-510

Sprinson JE, Lawton ES, Porco TC, Flood JM. and Westenhouse J.L (2006). Assessing the validity of tuberculosis surveillance data in California. *BMC Public Health*. 2006; 6:

Sutter RW, Kew, OM and Cochi SL (1988). Poliovirus vaccine-live. In: Vaccines, Fifth Edition. Eds. Stanley Plotkin, Walter Orenstein and Paul Offit. Saunders, Elsevier.

Thompson KM, Tebbens RJD, Pallansch MA, Kew OM, Sutter RW, Aylward RB, Watkins M, Gary HE, Alexander J, Jafari H, Cochi SL. (2008). The risks, costs and benefits of possible future global policies for managing polioviruses. *American Journal of Public Health* 98 (7): 1322-1330.

United Nations (2000). United Nations Millennium Declaration. Fifty-Fifth Session of the General Assembly. New York: United Nations; 2000. UN document A/RES/55/2. Available from http://www.un.org./millennium/ .

United Nations (2005). Investing in development. A practical plan to achieve the Millennium Development Goals. *United Nations Millennium Project*. Report to the UN Secretary General. London: UN Millennium Project.

Woodard S, Archer L, Zell E. Ronveaux O, Birmingham M, Design and simulation study of the immunization Data Quality Audit (DQA) (2007) . *Ann Epidemiol. 17(8):628-33.*

World Health Organization (1997). *Global Programme for Vaccine and Immunization, Expanded Programme on Immunization*. Geneva: World Health Organization. Available from: http://www.who.int/csr/ihr/polio1996en.pdf. Accessed on 24/04/2008.

World Health Organization (2000). *Global Eradication of Poliomyelitis by the Year 2000: resolution of the 41st World Health Assembly*. Geneva: World Health Organization,

World Health Organization Bulletin (2005). *The Immunization data quality audit: verifying the quality and consistency of immunization monitoring systems.* Bulletin World Health Organization 83 (7).

World Health Organization (2006). *The 56th Session of the WHO Regional Committee***,** Addis Ababa, September 2006. Available from: http://www.who.int/countries/eth/news/2006/rc56/en/index.html.

World Health Organization (2007). *The WHO Family of International Classifications*. Geneva: World Health Organization. Available from: http://www.who.int/health_topics/classification/en/.

World Health Organization (2007). *The Global Outbreak Alert and Response Network*, Geneva: WHO. Available from: http://www.who.int/csr/outbreaknetwork/en/ and http://www.who.int.health_topic/epidemiologic_surveillance/en/.

World Health Organization Bulletin (2009). Resurgence of wild polio type 1 and 3 in 15 African countries, January 2008-March 2009. Weekly epidemiological record. No.16 April 17, 2009 133-140.

World Health Organization Bulletin (2009). Progress towards interruption of wild polio transmission worldwide, 2009. Weekly epidemiological record. No.12 April 3, 2008, 308-12.

World Health Organization Bulletin (2009). Wild poliovirus type1 and type 3 importations – 15 countries, Africa, 2008-2009. Weekly epidemiological record. No.14 April 17, 2009 357-62.

World Health Organization Bulletin (2010). Progress towards interruption of wild polio transmission worldwide, 2010. Weekly epidemiological record. No.18 May 14, 2009 545-50.

World Health Organization Bulletin (2010). Update on vaccine-derived polioviruses—worldwide, January 2008-June 2009. Weekly epidemiological record. No.36 September 18, 2009; 1002-1006.

World Health Organization, (2001). Making surveillance work. Module 3: Logistics management. Department of Vaccines and Biologicals. WHO/IVB/01.10.

World Health Organization, (2003). WHO-Recommended standards for surveillance of selected vaccine-preventable diseases. Vaccines and Biologicals, World Health Organization, Geneva. WHO/V&B/03.01

World Health Organization, African Region (2005). *AFP Surveillance Field Guide*. Brazzaville, Congo. Pp 15

World Health Organization, African Region (2008). *AFRO – AFP surveillance Field Guide.* Brazzaville. World Health Organization.

World Health Organization, African Region (2008). *Polio Weekly Update.* Brazzaville. World Health Organization.

World Health Organization (2010). Outbreak of poliomyelitis in Tajikistan in 2010:risk of importation and impact on polio surveillance in Europe?. Euro Surveillance 15(17). pii: 19558

Wringe A, Fine PE, Sutter RW, Kew OM (2008). Estimating the extent of vaccine-derived poliovirus infection. PLoS ONE 3(10): e3433. Doi:10.1371/journal.pone.0003433

Yilmaz S., Hegedus J., and Bell M. (2003). *Sub-national data requirements for fiscal decentralization*. *Case study from Central and Eastern Europe*. Washington, DC: World Bank.

APPENDICES

Appendix I: Case Investigation Form

CASE INVESTIGATION FORM - - ACUTE FLACCID PARALYSIS
Ministry of Health

Official Use Only: EPID Number: _____-_____-_____-_____-_____ Received: ____/____/____
Country Region/Prov. Districts Year onset Case Number by the Programme at National level

IDENTIFICATION
District:_____ Region/Province: _____ nearest Health Facility: _____
Address:_____ Village:_____ City: _____
Father/Mother:_____
Patient name: _____ Age: _____years _____months
Date of Birth (DOB)___/_____/_____ (If DOB Unknown) Sex: ☐ M=Male F=Female

NOTIFICATION/INVESTIGATION:
Notified by:_____ Date of Notification___/_____/_____ Date of Investigation: ___/_____/_____

HOSPITALIZATION Hospitalized: ☐ 1=Y 2=N Date of admission to hospital, if applicable: _____/_____/_____
Hospital record #:_____ Name of hospital/Address:_____

CLINICAL HISTORY (1=Y, 2=N, 9=Unknown)
Fever at the onset of paralysis? ☐ Progressive Paralysis ≤3 days? ☐
Date of onset of paralysis: ___/_____/_____ Is Paralysis flaccid and acute? ☐ 1=Y, 2=N, 9=Unknown Asymmetric? ☐ 1=Y, 2=N, 9=Unknown
Site of Paralysis LA ☐☐ RA LL ☐☐ RL

AFTER INVESTIGATION, WAS THIS A TRUE AFP? ☐ 1=Y 2=N
If not, do not fill the rest of the form and record 6 under final classification

IMMUNIZATION HISTORY
Total Number of Polio vaccine doses ☐ Exclude dose at birth 99=Unknown
Doses at birth _____/_____/____ 2nd_____/_____/____ 4 TH_____/_____/____ If > 4
VPO 1st _____/_____/____ 3 rd_____/_____/____ Last _____/_____/____ dose

STOOL SPECIMEN COLLECTION:
_____/_____/_____ Date 1st specimen
_____/_____/_____ Date 2nd specimen
_____/_____/_____ Date specimen sent to the to the national level
_____/_____/_____ Date specimen received at the national level
_____/_____/_____ Date specimen sent inter-county/national Laboratory

STOOL SPECIMEN RESULTS:
___/_____/_____ Date combined Cell Culture Results available
___/_____/_____ Date Results sent to national EPI
___/_____/_____ Date Results received at national EPI
_____/_____/_____ Date specimen received at inter country (I-C)/national Lab
☐ 1= Adequate 2= Not adequate Status of specimen at Reception at the lab
Final cell Culture Results ☐ 1=Suspected poliovirus 2= Negative 3=NPENT 4-Suspect poliovirus + NPENT
_____/_____/_____ Date sent from I-C/National Laboratory to regional lab
_____/_____/_____ Date I-T differentiation results sent to EPI
_____/_____/_____ Date I-T differentiation results received at EPI
W1 W2 W3 ☐☐☐ 1=Y, 2=N
VDPV ☐ Type 1,2,3 4-Neg
V1 V2 V3 ☐☐☐ 1=Y, 2=N
(R) NPENT ☐ 1=positive, 2=Negative
NEV ☐

FOLLOW-UP EXAMINATION
___/_____/_____ Date of Follow-up exam
Residual Paralysis? LA ☐☐ RA LL ☐☐ RL
Results of exam ☐ 1= Residual Paralysis 2=No residual paralysis 3= Lost follow-up 4=Died before follow-up

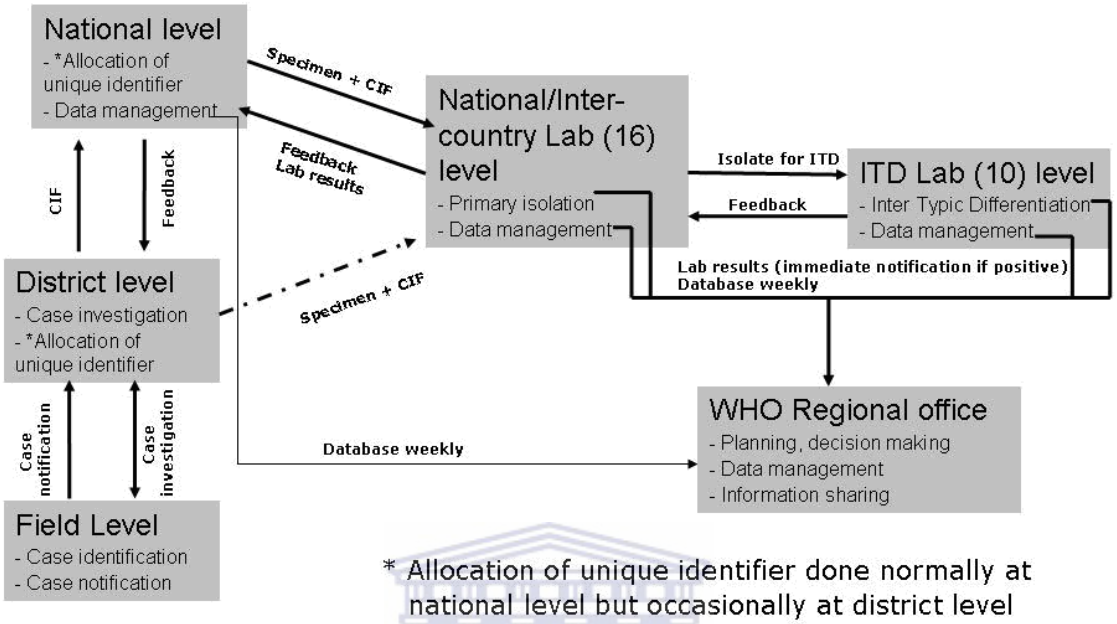
FINAL CLASSIFICATION ☐ 1=Confirmed Polio 2=compatible 3=Discard 6=Not an AFP case 7=cVDPV

INVESTIGATOR: Name_____ Title_____
Unit:_____ Address_____ Tel:_____

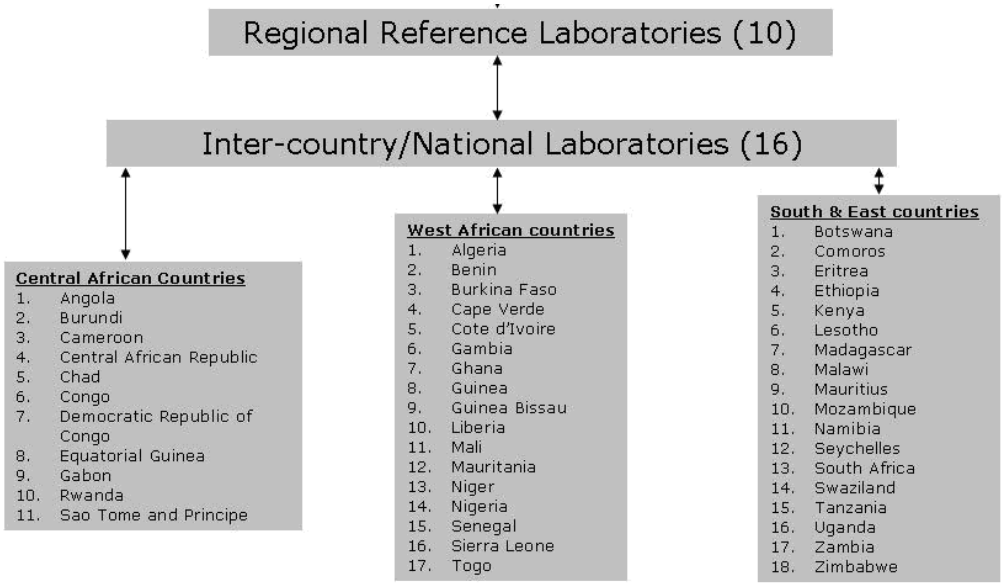
Appendix II: Program used to generate EPID errors

```
*************************************************************************
************************
* PGM for selecting cases with Epid numbers wrongly formulated
*************************************************************************
************************
ROUTEOUT 'Epidnumber_Issues' REPLACE
SORT Epidnumber
DEFINE CTRY
ASSIGN CTRY=SUBSTRING(EPIDNumber,1,3)

SELECT substring( EpidNumber , 4 , 1 ) <> "-" or substring( EpidNumber , 8 ,
1 ) <> "-" or substring( EpidNumber , 12 , 1 ) <> "-" or substring(
EpidNumber , 15 , 1 ) <> "-"
IF RecordCount>0 THEN
      TYPEOUT "Epidnumber Wrongly Formulated" (BOLD, UNDERLINE) TEXTFONT  4
      FREQ CTRY
      LIST EpidNumber NamesOfPAtient District DateOfOnset COLUMNSIZE=0
END
SELECT

SELECT Substring(Epidnumber,14,1)="x" OR Substring(Epidnumber,16,1)="x" OR
Substring(Epidnumber,14,1)="" OR Substring(Epidnumber,16,1)=""
IF RecordCount>0 THEN
      TYPEOUT "Epidnumber Incomplete" (BOLD, UNDERLINE) TEXTFONT  4
      FREQ CTRY
      LIST EpidNumber NamesOfPatient District DateOfOnset COLUMNSIZE=0
END

DEFINE YR_EPID
DEFINE YR_ONSET
ASSIGN YR_EPID=Substring(Epidnumber,13,2)
ASSIGN YR_ONSET=Substring(FORMAT(YEAR(DateOfOnset)),3,2)
SELECT NOT YR_EPID=YR_ONSET
IF RecordCount>0 THEN
      TYPEOUT "Epidnumber year and date of onset don't match" (BOLD,
UNDERLINE) TEXTFONT  4
      FREQ CTRY
      LIST EpidNumber NamesofPatient District DateOfOnset COLUMNSIZE=0
END
CLOSEOUT
```

Appendix III: Flow chart showing the collection, transfer and capture of both clinical and laboratory data, African Region



Appendix IV: Movement of specimens from field to laboratory by ISTs

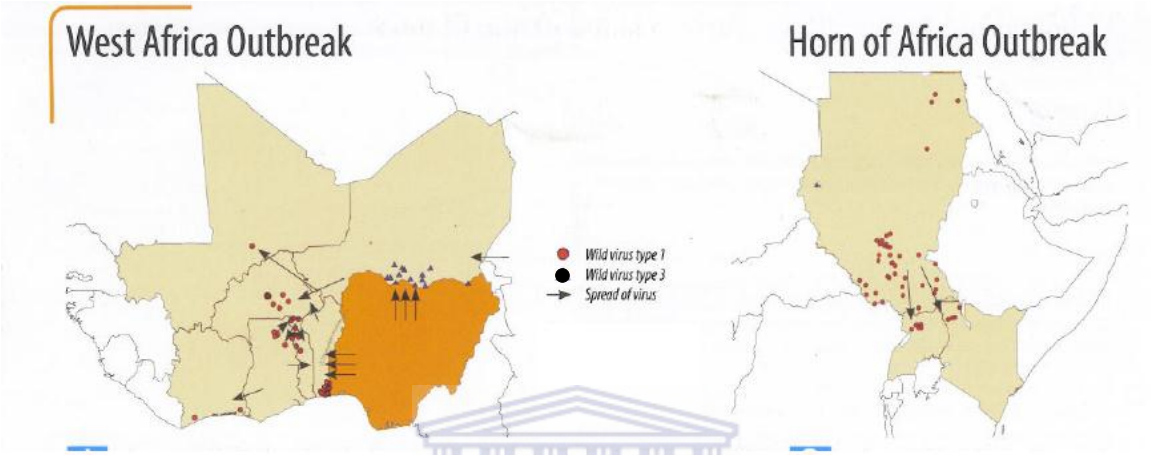


Appendix V: Assessment of various errors and their implications for polio eradication programme

Error type CODE	Identified errors in the clinical and laboratory databases and their frequency in %%	Importance of the error as per assessment scale			Possible Implications on PEP
		High	Medium	Low	
A1	Wrong formulation of EPIDNumber			X	- When selecting a such a case using EPI Info coding such a case may be missed
A2	Incomplete EPIDNumber	X			- Wrong location of case within or outside a country thereby initiating control measures where there is no virus
A3	Year of EPID Number and date of ONSET do not match	X			- Difficult to determine which year of onset of case is correct. It may affect the polio free Certification Commission decision as a country should be certified polio free if the last 3 years zero cases are reported.
B1	Cases missing date of onset	X			- Difficult to determine which year the case happened thereby interfering with proper AFP rate - May also affect stool adequacy in view of stool being collected within 14 days
B2	Cases with year of onset not matching with the year in the database		X		- Can increase or decrease the non-polio AFP rate for a particular year or location.
C1	Date of onset and date 1st stool collected with negative interval	X			- Difficult to determine how timeliness of specimen collected from date of notification
C2	Cases where date of onset and date 2nd stool was collected is greater than 14 days	X			- Reduces the chances of virus isolation - indicates on poor polio surveillance in the area
C3	Cases where the interval between date 1st stool was collected and date 2nd stool was collected is negative or greater than 48hr		X		- Affects chances for isolation of virus and in cases where 1st sample is inadequate or negative
D1	Cases missing the date specimen received in the lab	X			- difficult to determine the time spent in transportation of specimens - affects timeliness of specimen shipment
D2	Cases missing primary isolation results	X			- Delays in decision making more especially when a case is suspected poliovirus - delays in expert committee to classify cases
D3	Cases missing intra-typic differentiation results		X		- delays in decision making a mop up to undertake a mop up campaign especially when a case is confirmed polio virus - epidemiological linkages between AFP cases caused by same poliovirus type will be affected or impossible to establish

Based on the above assessment of errors and their implications for the GPEI in the African Region, hypotheses number 4, is therefore confirmed.

Appendix VI: Current status of wild poliovirus outbreaks in the West and Horn of Africa (data as of April 2009). The red and blue dots and triangles depict countries with circulating WPV types 1 and 3 respectively.



Source: Polionews No.32, 1st Quarter, 2009 EDITION. WHO/CDC/UN IEF (2009)