ROLE OF DATA QUALITY ASSESSMENT IN PERFORMANCE OF HIV PREVENTION PROJECTS IN NATIONAL AIDS AND STI CONTROL PROGRAMME, NYERI COUNTY, KENYA

BY

VALENTINE MBULWA NGUMBAU

A RESEARCH PROJECT REPORT SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF MASTER OF ARTS DEGREE IN PROJECT PLANNING AND MANAGEMENT, UNIVERSITY OF NAIROBI

2013
DECLARATION

This research report is my original work and has not been submitted in this university for the award of a degree or any other institution of higher learning for the award of any degree.

_____________________________  __________________________
Date

VALENTINE MBULWA NGUMBAU

L50/77511/2012

This research report has been presented for examination with my approval as the university supervisor.

_____________________________  __________________________
Date

Dr. Lillian Otieno-Omutoko

Senior Lecturer, Department of Extra mural studies

School of Continuing and Distance Education

University of Nairobi
DEDICATION

This research work is dedicated to my Father Stephen, Mother Joyce and my siblings for their unwavering support.
ACKNOWLEDGEMENT

I would like to thank the University of Nairobi for granting me the opportunity to pursue the Master of Arts degree in Project Planning and Management course. I would also like to thank the lecturers and staff of Nyeri Extra Mural Center for their support and providing a conducive environment for learning.

I sincerely thank my supervisor Dr. Lillian Otieno-Omutoko for her guidance, advice and encouragement and the patience she has shown in the entire process of writing this proposal. I thank her for her time and nurturing.

I also wish to thank my colleagues for their moral support and all the support staff of University of Nairobi, Nyeri Extra Mural Center, for the support they accorded me during the entire period I was pursuing the Masters of Arts Degree in Project Planning and Management.

I would like to thank the staff of the Nyeri NASCOP office for their encouragement, assistance and advice in the entire duration of this project and my colleagues for their guidance, patience and the information they were able to share with me as I undertake this project.

I wish to acknowledge and thank all the participants who took their time to complete and return the questionnaires and also took time to sit for interviews and made this research study possible.

I also thank my family and friends for providing moral support during this research period and all those who took time out of their way to enquire about my progress in the research.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>DECLARATION</td>
<td>ii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENT</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xi</td>
</tr>
<tr>
<td>ABBREVIATIONS AND ACRONYMS</td>
<td>xii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>xiii</td>
</tr>
<tr>
<td>CHAPTER ONE</td>
<td>1</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Background Information</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Problem Statement</td>
<td>7</td>
</tr>
<tr>
<td>1.3 Purpose of the Study</td>
<td>8</td>
</tr>
<tr>
<td>1.4 Objectives of the Study</td>
<td>8</td>
</tr>
<tr>
<td>1.5 Research Questions</td>
<td>8</td>
</tr>
<tr>
<td>1.6 Significance of the Study</td>
<td>8</td>
</tr>
<tr>
<td>1.7 Scope of the Study</td>
<td>9</td>
</tr>
<tr>
<td>1.8 Assumptions of the Study</td>
<td>9</td>
</tr>
<tr>
<td>1.9 Limitations of the Study</td>
<td>9</td>
</tr>
<tr>
<td>1.10 Definition of Significant Terms</td>
<td>10</td>
</tr>
<tr>
<td>CHAPTER TWO</td>
<td>12</td>
</tr>
<tr>
<td>LITERATURE REVIEW</td>
<td>12</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>12</td>
</tr>
<tr>
<td>2.2 Historical Perspective of Data Quality and Data Quality Assessment</td>
<td>12</td>
</tr>
<tr>
<td>2.3 Assessment of Data Quality</td>
<td>14</td>
</tr>
<tr>
<td>2.4 Data Quality Assessment as an Informative Process</td>
<td>15</td>
</tr>
<tr>
<td>2.4.1 Data Management System</td>
<td>15</td>
</tr>
<tr>
<td>2.4.2 Dissemination of Data Quality Assessment Information</td>
<td>17</td>
</tr>
<tr>
<td>2.5.1 Data Quality Assessment Standards and Procedures</td>
<td>20</td>
</tr>
</tbody>
</table>
2.6 Data Quality Assessment and Data Quality Improvement ............................................. 25
2.6.1 Plan-Do-Check-Act Cycle and Process Improvement ............................................. 25
2.6.2 Improvement Methods in Data Quality Assessment ................................................. 27
2.7 Conceptual Framework ......................................................................................... 29
2.8 Conclusion ....................................................................................................... 30

CHAPTER THREE .................................................................................................... 31
RESEARCH METHODOLOGY ..................................................................................... 31
3.1 Introduction ........................................................................................................ 31
3.2 Research Design ............................................................................................... 31
3.3 Target Population .............................................................................................. 31
3.4 Sampling Procedure and Sample Size ............................................................... 32
3.4.1 Sampling Procedure ...................................................................................... 32
3.4.2 Sample Size ................................................................................................ 32
3.5 Research Instruments ....................................................................................... 33
3.6 Pilot Study ........................................................................................................ 34
3.6.1 Validity of Research Instruments ................................................................... 34
3.6.2 Reliability of Research Instruments .............................................................. 35
3.7 Data Collection Procedure ................................................................................ 35
3.8 Data Analysis .................................................................................................... 36
3.9 Ethical Considerations ....................................................................................... 36
3.10 Operationalization Table ................................................................................ 37
3.11 Dissemination Plan .......................................................................................... 38
3.12 Summary ........................................................................................................ 38

CHAPTER FOUR .................................................................................................... 39
DATA ANALYSIS, PRESENTATION AND INTERPRETATION ........................................ 39
4.1 Introduction ........................................................................................................ 39
4.2 Response Rate .................................................................................................. 39
4.3 Background Information of the Data clerks, District Aids and STI Co-ordinator and Health Records Information Officers ................................................. 40
4.3.1 Gender of the Respondents .......................................................................... 40
4.3.2 Age of the Respondents ............................................................................... 41
4.3.3 Level of Education of Respondents ................................................................. 42
4.3.3 Positions Held ........................................................................................................ 44
4.4 Data Collection and the Informative Process .......................................................... 45
4.4.1 Frequency of collecting Project Related Data ...................................................... 45
4.4.2 Project Website ...................................................................................................... 46
4.4.3 Rating of Aspects Related to Data Collection and the Informative Process ........ 46
4.4.4 Ways Used to Ensure the Findings are known ...................................................... 48
4.4.5 Ways Project Management Teams can help in ensuring the Data Collection and Information Dissemination is more effective .................................................. 49
4.4.6 Average Duration of the Projects Undertaken .................................................... 49
4.4.7 Critical Issues from the Data Collected ............................................................... 50
4.4.8 Important components of quality Data that are assessed in the Projects ............. 51
4.4.9 Management of Expectations of Data Quality from Various Audiences .............. 51
4.4.10 Most Common Data Problems Encountered ...................................................... 51
4.4.11 Frequency of Disseminating Information from the Assessment ......................... 51
4.4.13 Common Ways Used to Disseminate Information .............................................. 52
4.4.14 Procedures Used to Facilitate the Process of Dissemination ............................. 53
4.5 Standard Operating Procedures ............................................................................. 53
4.5.2 Rating of Aspects Related to Standard Operating Procedure ............................... 55
4.5.3 Frequency of Performing Data Quality Assessment ............................................. 56
4.5.4 Skills the Staff Performing Data Quality Assessment need have to perform the Process .......................................................... 57
4.5.5 Suggestions to Help in Making the Process more Effective and Efficient .......... 57
4.5.6 If the Data Quality Assessment Process made the Impact in the Overall Performance of the Project ......................................................................................................... 57
4.5.7 Presence of Checklist to Facilitate the Data Quality Assessment Process ............ 58
4.6 Data Quality Improvement ....................................................................................... 59
4.6.1 Training as an Improvement Mechanism .............................................................. 59
4.6.2 Rating of Aspects Related to Data Quality Improvement ........................................ 60
4.6.4 Measures and Steps taken to Ensure that there Continuous Improvements ........ 61
4.6.5 Methods that Work Best in Improving the Process of Data Quality Assessment ...... 62
4.7 Summary .................................................................................................................. 62
CHAPTER FIVE ................................................................................................................................................. 63
SUMMARY OF FINDINGS, DISCUSSION, CONCLUSIONS, RECOMMENDATIONS AND AREAS FOR FURTHER RESEARCH .......................................................................................................................... 63
5.1 Introduction .................................................................................................................................................. 63
5.2 Summary of Findings .................................................................................................................................. 63
5.2.1 Data collection and the Informative Process of Data Quality Assessment in Performance of HIV Prevention Projects .................................................................................................................. 63
5.2.2 Standard Operating Procedures of Data Quality Assessment in Performance of HIV Prevention Projects ........................................................................................................................................ 63
5.3.3 Data Quality Improvement of Data Quality Assessment in Performance of HIV Prevention Projects ........................................................................................................................................ 64
5.3 Discussion of Findings ................................................................................................................................. 64
5.4 Conclusions ................................................................................................................................................ 65
5.5 Recommendations ..................................................................................................................................... 66
5.6 Areas of Further Research .......................................................................................................................... 66
5.7 Summary .................................................................................................................................................... 67
REFERENCES .................................................................................................................................................. 68
APPENDICES .................................................................................................................................................. 72
Appendix 1: Letter of Transmittal .................................................................................................................. 72
Appendix 2: Letter of Authority from National Commission for Science and Technology ............... 73
Appendix 3: Letter of Authority from Nyeri National AIDS and STI Control Program Office ....................................................................................................................................................... 74
Appendix 4: Letter of Introduction from University ...................................................................................... 75
Appendix 6: Interview guide for Project Managers and Data Analysts ....................................................... 80
Appendix 9: Data Quality Assessment Check list ............................................................................................ 86
LIST OF TABLES

Table 3.1: Population Size ................................................................................................................. 32
Table 3.2: Operationalization Table ........................................................................................................ 37
Table 3.3: Dissemination Plan ........................................................................................................................... Error! Bookmark not defined.
Table 4.1: Response Rate ........................................................................................................................... 39
Table 4.2: Gender of Data Clerks, District Aids and STI Co-coordinators and Health Records Information Officers ................................................................................................................................................................................. 41
Table 4.3: Gender of Project Managers and Data Analysts ................................................................. 41
Table 4.4: Age of Data Clerks, District Aids and STI Co-coordinators and Health Records Information Officers ................................................................................................................................................................................. 42
Table 4.5: Age of the Project Managers and Data Analysts ........................................................................ 43
Table 4.6: Level of Education of Data clerks, District Aids and STI Co-coordinators and Health Records Information Officers ................................................................................................................................................................................. 44
Table 4.7: Level of Education of Project Managers and Data Analysts ................................................ 44
Table 4.8: Positions Held by Data clerks, District Aids and STI Co-coordinators and Health Records Information Officers ................................................................................................................................................................................. 44
Table 4.9: Positions Held by Project managers and data analysts (n=5) .................................................. 44
Table 4.10: Frequency of collecting Project Related Data .......................................................................... 45
Table 4.11: Project Website ......................................................................................................................... 46
Table 4.12: Rating of Aspects Related to Data Collection and the Informative Process ...................... 47
Table 4.13: Average Duration of the Projects Undertaken ......................................................................... 50
Table 4.14: Frequency of Disseminating Information from the Assessment .......................................... 52
Table 4.15: Procedures Used to Facilitate the Process of Dissemination .............................................. 53
Table 4.16: Systematic Procedures that Govern the Process of Data Collection according to Data clerks, District Aids and STI Co-coordinators and Health Records Information Officers .................................................. 54
Table 4.17: Systematic Procedures that Govern the Process of Data Collection according to Project managers and Data Analysts ................................................................................................................................................................................. 54
Table 4.19: Frequency of Performing Data Quality Assessment .................................................. 56
Table 4.20: Skills the Staff Performing Data Quality Assessment need have to perform the Process .................................................................................................................................................. 57
Table 4.21: Presence of Checklist to Facilitate the Data Quality Assessment Process ............. 58
Table 4.22: Training as an Improvement tool .................................................................................. 59
Table 4.23: Rating of Aspects Related to Improvement ................................................................ 60
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Components of a Data Management System</td>
<td>16</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Flowchart for Universal Data Quality Assessment</td>
<td>22</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Plan-Do-Check-Act Cycle</td>
<td>25</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Implementation of Improvement Process</td>
<td>27</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Conceptual Framework</td>
<td>29</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------</td>
<td></td>
</tr>
<tr>
<td>CDC</td>
<td>Center for Disease Control</td>
<td></td>
</tr>
<tr>
<td>DQ</td>
<td>Data Quality</td>
<td></td>
</tr>
<tr>
<td>DQA</td>
<td>Data Quality Assessment</td>
<td></td>
</tr>
<tr>
<td>HIV</td>
<td>Human Immunodeficiency Virus</td>
<td></td>
</tr>
<tr>
<td>IMF</td>
<td>International Monetary Fund</td>
<td></td>
</tr>
<tr>
<td>NASCOP</td>
<td>National Aids and STI Control Program</td>
<td></td>
</tr>
<tr>
<td>NHRL</td>
<td>National HIV Reference Library</td>
<td></td>
</tr>
<tr>
<td>PEPFAR</td>
<td>President’s Emergency Plan for AIDS Relief</td>
<td></td>
</tr>
<tr>
<td>SOP</td>
<td>Standard Operating Procedures</td>
<td></td>
</tr>
<tr>
<td>USAID</td>
<td>United States Agency for International Development</td>
<td></td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
<td></td>
</tr>
</tbody>
</table>
ABSTRACT

HIV and Aids has been a global pandemic since the first cases were discovered in the early 1980’s. Statistics from the Foundation for AIDS Research estimate that 34 million people are living with the virus worldwide with an estimated 2.5 million new infections every year as at 2011. In Kenya, as at December 2011, 1.6 million people were living with the virus. These statistics especially those on new infections, have led to prevention projects that aim to educate the population on how to prevent the spread of the virus and in the forefront of these efforts is the National AIDS and STI Control Program which was established in 1987 to help the government in the fight against HIV. The objectives of this research was to investigate how the data collection and informative process of DQA works to influence performance of these projects, to identify how the operating procedures of DQA influence project performance and establish how data quality improvement in DQA influences performance of these projects. A descriptive survey research design was used in this study and from the population selected; stratified random sampling as well as purposive sampling was used to choose the desired sample. Data was collected through the use of questionnaires and interview guides and data collected was analyzed using The Statistical Package for Social Sciences program. The findings of the study showed that data collection procedures and tools as well as dissemination of DQA information played a role in performance of projects. It also found that there were standard operating procedures in the projects under the study that made the processes efficient and effective and also found data quality improvement can be worked on as the projects evolve. The study revealed that data collection tools and procedures play a major role in the quality of data produced and that the dissemination of the findings of DQA will improve the quality of data. It also revealed that standard operating procedures need to be present in the DQA process and how data quality improvement provides a quality spectrum in the DQA process. The study concluded that there was more that needed to be done to enhance DQA such as adopting technology and follow up in the usage of the procedures as well as providing the staff dealing with DQA a more targeted training so as to improve their skills which will in turn lead to an overall improvement of data quality. It recommended that there should be project websites to track progress and help in dissemination as well as having targeted training based on the job description.
CHAPTER ONE

INTRODUCTION

1.1 Background Information

The first cases of HIV in the world were discovered in 1981 with the first case in Kenya discovered in 1984 and since then the numbers have kept on rising. Statistics from the Foundation for AIDS Research estimate that 34 million people are currently living with HIV/AIDS with a further estimated 2.5 million new infections. In Kenya, as at December 2011, 1.6 million were living with the virus. After the first cases of HIV were discovered, the Kenya National Aids and STI Control Programme (NASCOP) was established in 1987 to spearhead the Ministry of Health interventions on the fight against HIV/AIDS. It has several interventions and projects whose aim is to fulfill its mission which is to prevent HIV/STI infections and improve quality of life for those infected and affected by providing strategic information, developing guidelines, policies, co-ordination of partners, provision of quality HIV/STI services in order to mitigate the socioeconomic impact of HIV/STI in Kenya (nascop.or.ke).

The prevention projects are designed to help control the spread of HIV and include those targeting the Most at Risk Populations, Injecting Drug Users (IDU’s), discordant couples, and the commercial sex workers. They facilitate the distribution of condoms, ARV’s, counseling and seminars that address all aspects of the spread and control of the virus. The introduction of the prevention projects has seen the number of new infections fall. The number of new HIV infections among adults in 2010 was less than one-third the numbers reported in 1993, when the country’s epidemic peaked. An estimated 49,126 people died of AIDS-related causes in 2011, slightly more than one-third the annual numbers who died in 2002–2004.

National Programs and donor-funded projects, which form the bulk of most NASCOP projects, are working towards achieving the set out goals in the fight against HIV. Measuring the success and improving the management of these projects is predicated on strong Monitoring and Evaluation (M and E) systems that produce quality data related to program implementation. It is with this reason that the production of quality data is an integral part in checking for the success of these projects which in turn help in providing an accurate basis for decision making and the acquisition of funds from the partners supporting these projects. The data that is usually collected
in these projects includes data that shows the number of infected persons, number of new infections, those under ARV drugs, discordant couples, most at risk populations, the different ARV regimens and how they are working, adverse drug reaction information and many more. This information is used to allocate funds to these projects, note the trends that are appearing and where to concentrate more efforts. It also helps in noting which interventions are working and which ones are not. The data is collected through registers that are developed and distributed to the districts and facilities. With such important data, it has become very important to ensure that the data produced from these projects is of utmost quality and reliable.

Data Quality Assessment (DQA) refers to the scientific evaluation of environmental data to determine if it meets the planning objectives of the project and is therefore of the right type, quality and quantity to support its intended use (EPA QA/G9). The term environmental data was used by the EPA to refer to data about the environment which was the main focus of their research but over time this term has been used to refer to data representing the area where it is collected and used. In this case the environment will be that of the HIV prevention projects. Production of high quality data depends on assessment of data which involves ensuring that the quality indicators specified before the onset of the project are present in the data collected. Many projects have therefore introduced the use of a data quality mechanism to check for any data quality problems that may be present in the collected data and also helps in providing a foundation for informing the users on how that data should be used. It was introduced as a way of assessing the quality of data produced and to provide checks and ensure the data produced is meaningful. Without a data quality assessment mechanism, most projects face the risk of losing control of the various processes from data collection to data analysis and presentation. According to UNAIDS (2008) the objectives of a data quality assessment mechanism in any project is to verify the quality of reported data for key indicators at selected sites and to assess the ability of data management systems to collect and report quality data.

The European Statistics Code of Practice (2011), data quality assessment is important in analyzing the data quality components such as relevance and accuracy and also in ensuring there is a systematic procedure and assessment of the processes involved. In recent years, most European countries have had efforts in implementation of data quality assessment methods in their projects using methods such as self-assessment, auditing and user satisfaction surveys.
Data Quality Assessment mainly primarily focuses on verifying the quality of reported data, and assessing the underlying data management and reporting systems for standard program-level output indicators. According to David Boone (2008) DQA is not intended to assess the entire M&E system of a projects response to HIV/AIDS. In this context, the DQA relates to components of Supportive supervision, which aims to ensure that the objectives set out are achieved as well as helping in the managing of the process improvement and the needs of all the stakeholders are met, and data auditing of the program which aims in providing an assessment of data quality and M&E systems in selected projects.

Data of high quality has six key attributes and the degree to which data meets these attributes determines its quality. The attributes include timeliness, completeness, appropriateness, accuracy, reliability and credibility. Different countries use different methods of ensuring data quality in their projects. In Canada, they have developed the Quality Assurance Framework which is a set of management, operating and consultative practices, procedures and mechanisms. The framework is used by statistics Canada to help manage the quality of information arising and produced in the lifetime of these projects. It has been adapted and it links the projects’ overall objectives and provides clear avenues for feedback, performance review and ongoing planning. Their assessment of the data is done using the context of the six elements of data quality which they describe to be relevance, accuracy, timeliness, accessibility, interpretability and coherence.

According to the framework stipulated and to ensure proper management of these six dimensions, it is the duty of the project leaders to create an environment and develop a culture that recognizes the importance of data quality and the role it plays in ensuring the project is effective and efficient in meeting the overall objectives. Data quality is important because it helps in managing accountability, ensuring effective of service, helps in prioritizing the best use of resources in the project and also as a check for achievement in the project. Projects that are data intensive, such as the HIV prevention projects, the data generated should follow the life cycle of data (Batini and Scannapieco 2006). The life cycle entails the three steps of planning, implementation and assessment. DQA completes the lifecycle by providing the evaluation needed to determine if performance of the project is as expected. The steps in conducting the
assessment but present a down side in that the process cannot conclusively prove the project objectives have been met due to the use of a sample

The project implementers should include verified programs of data analysis and research to encourage continuous improvement within the project environment (Statistics Canada, 2000d). The Manitoba Center for Health Analysis in most of their health projects, including those dealing with HIV, has incorporated methods such as identifying outliers, which involves statistical approaches, to access the quality of data. They have also embraced automation which links several databases to a main frame and having a data quality website. Despite rolling out these measures, there have been several limitations such as standardization of the data collected and having a central format library. In the United States, the Environmental Protection Agency (EPA) developed a Data Quality Assessment Guide that has been used as the reference guide in most statistical projects and it helps and guides project managers and planners determine the type, quantity and quality of data needed to support the project decisions and ensure objectives are met. The guide has a clear five step process for performing DQA that are iterative and continuous.

The first step is to review the projects objectives and sampling design where the objectives defined during systematic planning to assure that they are still applicable. If objectives have not been developed (e.g. when using existing data independently collected), specify them before evaluating the data for the projects objectives. Review the sampling design and data collection documentation for consistency with the project objectives observing any potential discrepancies. Conducting a preliminary data review follows where review reports for the validation of data, calculate basic statistics, and generate graphs of the data. Use this information to learn about the structure of the data and identify patterns, relationships, or potential anomalies. The next step is to select a statistical method and it entails selecting the appropriate procedures for summarizing and analyzing the data, based on the review of the performance and acceptance. The final steps are verification of the assumptions of the statistical method and drawing final conclusions from the data.

The steps outlined by the DQA have been used as a benchmark that other projects not statistical in nature and organizations have borrowed heavily on to come up with more suitable steps to suit their needs. For most health related projects the steps of conducting DQA have been streamlined
with the steps varying from one environment to another. The commonly used procedure as specified by United States Agency for International Development, Center for Disease Control and Global Fund has the steps of first determining the purpose of the assessment and selecting the level of assessment and the areas to be included in the assessment. Indicators that are most important to the assessment are identified as well as the data sources to be used as well as the reporting period. The next step is conducting the actual site visit and conducting the assessment. The results and findings are reviewed and finally a system strengthening plan for the areas found unsatisfactory is developed as well as follow up plans on where there was more that needed to be done. From the above steps the role of the DQA process as a tool to facilitate improvement comes out clearly with the system strengthening plan for the rest of the project as well as future projects.

In Africa, the USAID Assessment Report of May 12, 2010, indicated that Data Quality in HIV projects has been emphasized with interventions on DQA being undertaken in Botswana which has the second largest HIV prevalence rate in the world according to USAID and in South Africa where the WHO undertook an intervention in projects done at KwaZulu-Natal Province. It was established that accurate and reliable information is essential for monitoring and evaluating as well as improving the services provided in all the projects undertaken. The data quality improvement intervention in South Africa took place between May and November 2008. Without active monitoring through regular site visits and data verification, USAID projects in Botswana did not always have reasonable assurance that data used for performance-based decision making and for reporting were valid and reliable. A program to monitor data quality through regular site visits identified documentation and reporting issues and prevented many of the problems with data reliability identified in this report. Most African countries in their HIV projects tend to use the DQA as per the USAID and PEPFAR guidelines since this is the main source of funding for these projects and the CDC in collaboration with other agencies produced a document as a referral for quality HIV related data.

In Kenya, and more so NASCOP, has adopted the CDC, Global Fund and USAID protocols for the DQA. The management, at all levels so as to ensure consistent data quality, conduct an internal DQA and an independent body conducts external DQA to allow unbiased feedback on the quality of data. The parameters of quality data are validity, reliability, timeliness, integrity
and precision although these may vary depending on the environment they are being used and the type of data required. The DQA plan specifies the measures that must be taken of the data management process to ensure that each of the data quality parameters are maintained, and thus guarantee the quality of reported and used data. Each of the data quality parameters must be managed at each point in the data life cycle and there must be a verifiable trail that indicates that the data maintains its quality at each point in the data life cycle. In order to monitor and measure the quality of data at each point in the data management cycle, data quality assessments should be conducted regularly as the costs of low quality data impact at every level of project implementation.

The DQA can be conducted bi-annually, to allow for any recommendations to be implemented or annually, depending on the need and the objectives of the assessment. Quality assessment is meant to motivate both the client and the staff and, most of all, improve service delivery. Although the benefits of high-quality data may be readily apparent, significant programmatic risks are involved in reporting or using poor-quality data for decision making. The importance of data quality can be summarized using this quote by Bill Gates “Virtually everything in business today is a differentiated commodity, except how a company manages its information; how you manage your information determines whether you win or lose” (p. 5). This goes to show that data and especially quality data may be the determining factor of whether an Endeavour will be successful or a failure.
1.2 Problem Statement

The process of DQA, its application and use have been researched over time. A study done by Pippino (1999) discusses the aspects that the DQA process looks at in ensuring a well done process. It looks at the components that are key for a successful process. The study looks at the process in totality and also discusses how the process can be changed to suit different needs. The study shows a step by step account of the process is carried out and the people involved at every stage. In as much as this study provided insight on a process that is little known, it has various shortcomings in that it has no foundation. The researcher goes directly to the process and there is no discussion on why organizations should use this process and the role it plays in improving the output. A research done by Jones (2003) looks at how data quality can be improved through analysis of variables. The study reports that DQA can be validated by aspects such as quality reports and indicators. The research mainly looks at the output of the DQA process and ways of improving the overall process and achieving maximum results from the process. The output of the research was well received especially due to the improvement strategies involved. The study, just like the one done by Pippino assumes that the reason for undertaking DQA is already known and does not discuss why the process is important. A study done by statistical organizations in various countries such as EPA and Canadian Institute lean towards coming up with a methodological approach and discuss DQA in more of a statistical and procedural manner. A more detailed research undertaken by the WHO on DQA mostly deals with it as a process, how it will be adopted in various projects and has also leaned towards a methodological approach.

The research cited above show the leaning of the researchers towards DQA as a process, its application and methodologies used as well as its components. However, there is a gap left since they do not discuss the functions and roles of the process in the various organizations and projects they cite in their studies. In as much as they have given insight on the process itself, some organizations and project managers find it difficult to incorporate it since it is not clear why the process is important. This study therefore seeks to show the role of DQA in performance of these projects, how the process can improve performance and perhaps change the perception that it does not add value and eventually many more projects will incorporate it. According to David Boone (2008) DQA is not intended to assess the entire M&E system of a country’s response to HIV/AIDS. In this context, the DQA relates to components of Supportive supervision and data auditing of the program.
1.3 Purpose of the Study
The purpose of the study was to assess the role of data quality assessment in enhancing performance of HIV projects undertaken by NASCOP.

1.4 Objectives of the Study
This study was guided by the following objectives:-

1. To assess how data collection and the informative process in DQA influences performance of HIV projects in NASCOP.

2. To establish how Standard Operating Procedures in DQA influence performance of HIV projects in NASCOP.

3. To establish how data quality improvement in DQA influences performance of HIV projects in NASCOP.

1.5 Research Questions
1. How does data collection and the informative process of DQA influence performance of HIV projects in NASCOP?

2. How do the Standard Operating procedures on DQA influence the performance of HIV projects in NASCOP?

3. How does data quality improvement in DQA help in improving performance of HIV projects in NASCOP?

1.6 Significance of the Study
This study is significant to the data clerks and data management assistants because it puts into perspective effects of poor data quality in the overall project performance and achievement of set targets. It will also show appropriate data practices that can be employed to improve performance and also help the relevant stakeholders in making informed decisions and assist in providing justification for funding.

The findings of this study will also be of benefit to various project stakeholders who include the donors, the various department heads and the partners that help in these projects. It will help the data collection tools development team, tasked with developing the data collection tools, to come up with tools that capture what is relevant to the project and also develop better project indicators.
to track performance. It will also benefit the project managers to factor in, budget and prepare for the DQA process in the course of the projects life cycle which is important when performing a cost benefit analysis as regards DQA as well as justify the need for personnel and funds.

It is significant to other researchers interested in the realm of data quality as it will provide a reference in the area of data quality assessment as well as providing suggestions for further research.

1.7 Scope of the Study
The study will be carried out within the NASCOP HIV prevention projects dealing with the Most at risk populations (MARPS), HIV testing and counselling undertaken in Nyeri County in Nyeri Town Center, Mathira, Othaya, Kieni and Mukurweini.

1.8 Assumptions of the Study
The study assumed that the respondents will be truthful in the data they give and that it will represent the current situation.

1.9 Limitations of the Study
The research is limited due to the fact that the findings may not be generalized to all HIV prevention projects in general.
1.10 Definition of Significant Terms

This section aims at providing an operational definition to some of the commonly used terms and concepts used in this study.

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessment</td>
<td>Ways of having checks to ensure the data is correct from collection to analysis and final use.</td>
</tr>
<tr>
<td>Data</td>
<td>This is the collected information that is not processed.</td>
</tr>
<tr>
<td>Data Quality</td>
<td>Data that is accurate enough to be used for its intended purpose without bringing wrong conclusions; it is fit for its intended use in operations, decision making and planning.</td>
</tr>
<tr>
<td>Data Quality Assessment</td>
<td>Is a process of verifying the quality of data, assessing the system that produces that data from data collection to analysis and developing actions to improve both the process of collection, analysis and assessment as well as the system used for assessment.</td>
</tr>
<tr>
<td>Environmental data</td>
<td>Data that represents the area where it is collected.</td>
</tr>
<tr>
<td>Informative Process</td>
<td>A process that ensures information is available at the right place, right time and to the right people.</td>
</tr>
<tr>
<td>Performance</td>
<td>This describes the overall achievement of the project in relation to the objectives set out. Performance may be either negative or positive that is the objectives may be achieved or not.</td>
</tr>
<tr>
<td>Prevention projects</td>
<td>These are projects undertaken to help in the prevention of the spread of HIV.</td>
</tr>
<tr>
<td><strong>Data Quality Reliability</strong></td>
<td>Refers to a state of ensuring the data collection and assessment processes are stable and consistent over time.</td>
</tr>
<tr>
<td>----------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Standard Operating Procedures</strong></td>
<td>Refers to the guidelines stipulated to be followed when undertaking any particular process.</td>
</tr>
<tr>
<td><strong>Data Quality Validity</strong></td>
<td>Refers to the system of checking and ensuring the data is adequately representative of the performance.</td>
</tr>
</tbody>
</table>
CHAPTER TWO
LITERATURE REVIEW

2.1 Introduction
In this chapter, the researcher focuses on the role of Data Quality Assessment in improving performance of HIV projects and will be guided by the research objectives and research questions. It will discuss the historical perspective of data quality and DQA, literature related to DQA and several related aspects. This chapter will also include the conceptual framework of the study, suggestions for further research and the conclusion.

2.2 Historical Perspective of Data Quality and Data Quality Assessment
In the data quality realm, it has often been seen that data quality is not about technology but more about people. The data quality issue has been there since the introduction of the relational database (Sarsfield, 1999). Data quality assessment can be traced back to more than a century ago in 1848-1853 in Belgium while workers were tracing the biological indicators of pollution (Liebmann, 1962). The project was data intensive and covered large areas and there was no specific framework to be used in the assessment of this data. It was a time consuming endeavor but the information from the assessment was seen to greatly improve the efforts of the project.

In USA, another major data quality assessment intervention was underway throughout the country as they tried to improve the quality of healthcare between as early as 1910 and was continuous till the 1980’s (Donabedian, 1966). It involved hospitals and healthcare facilities throughout the country and the results of this assessment were used to recommend and develop guidelines and standards used in the health sector. In most European countries, data quality came with its own set of problems. They emphasized the inclusion of the basic dimensions of data (Ivanov, 1972). It is with these problems that systematic ways and procedures of carrying out DQA were developed. The United States EPA developed the DQA guide; adopted mainly for statistical projects, with the first draft in the 1990’s and consequent reviews being done the latest in 2006. The document aimed to provide guidance to organizations on DQA and provide specifications for proper decision making. The process was developed for project managers and planners to ensure the type, quantity and quality of data needed to achieve the objectives and aid in decision making is met. Frameworks to guide the DQA process were also developed and they included the “Zero defect data” framework which was mainly adopted for statistical processes to
control data quality (Hansen, 1991). Statistical organizations in America and other European countries e.g. Eurostat, after implementation of data quality procedures, came up with checks to ensure the final data product can be used as intended and this was a step towards the evolution of the DQA process. In 1993, United States president Clinton signed the Government Performance and Results Act (GPRA) where institutions were required to develop plans on how they will deliver quality products and services. This act was then used to appeal to institutions that dealt with data to embrace the quality agenda.

In Africa, DQA has been incorporated for as long as donor projects have been around. In 1997, South Africa, Ethiopia and Malawi, the UNESCO Institute of Statistics introduced the use of DQA in the population surveys and ensured use of standardized definitions was adopted (Stats, 1998). From as far as 1993, oil and gas producing countries in West Africa also use DQA as a way of monitoring the supplies and an important monitoring and evaluation tool. The uptake of the process has been on the rise in almost all major sectors that drive the economy. African Virtual University developed a DQA framework that borrowed from the international best practices where data from different institutions of higher learning was collected and the framework was adopted in 2007. Another sector that has embraced DQA is that of health where donors such as USAID and Global fund have stipulated as a requirement that the process be incorporated since it is seen as an accountability tool. Also in South Africa, Statistics South Africa which is the agency responsible for collection and dissemination of statistics developed their own assessment framework in 2008 showing that organizations are embracing the process.

In Kenya the subscription to the General Data Dissemination System (GDDS) on October 29, 2002 required a DQA performed on various institutions including the Central Bank, Ministry of Finance among others. The team chosen applied the IMF DQA Framework with specific indicators (IMF, 2005). There has not been a specific framework developed in the country and most used those developed by the EPA, IMF and USAID. Management Systems International (MSI) conducted a DQA involving 30 indicators for the USAID/Kenya Mission. It spanned across five major sectors which were youth development, economy, environment, resource management and education in 2012. It was to enable a thorough understanding of strength and weaknesses of data used in these areas and what can be done to improve it. Also in 2102, RTI
International had calls for the development of a QA protocol for neglected tropical diseases and it is through such endeavors that we hope to see a local framework.

Sound policies, allocation of resources and overall day to day management decisions in the projects being undertaken require timely and quality data. In the past, the systems employed in most of these projects have been seen as being inadequate as far as production of quality data is concerned (Hotchkiss, et al, 2010). High quality data must have a positive and continuous effect on any projects/programs measurements for outcome. It is these measurements that are then monitored as a means for improving project performance.

2.3 Assessment of Data Quality

Data quality assessment has been seen as an important tool for improving performance (Ghosh, 2011). The aspects that DQA looks at have been refined to ensure that the data and the data collection process both achieve the standards set to ensure performance. These aspects include validity and reliability. Whether there is a solid, logical relation between the project being undertaken and what is being measured. It looks at whether the data adequately represents performance of the project and analyses the data collection instruments to ensure they do not produce biased data. It includes aspects of measurement which check on whether the data collection instruments are well designed, whether the definitions for data to be collected operationally precise and also the training levels of the data collection team. The reliability aspect looks to ensure that there are stable and consistent data collection processes and analysis methods. It looks at the consistency of the processes, ensures that there is a clear laid down procedure that govern the process and also looks at transparency to ensure the procedures are well documented and errors are reported.

Another aspect is timeliness and this has to do with checking the frequency with which data is collected and the currency of the data being analyzed. The precision of this process is also important as it ensures that only an acceptable margin of error is allowed. The final aspect is integrity which ensures that the data is free of manipulation. All the above aspects serve to ensure that the performance of the projects based on the data being assessed is excellent leading to the achievement of objectives.
2.4 Data Quality Assessment as an Informative Process

The informative process is seen as a descriptive process that shows in details all pertinent information. It ensures all information required is available at the right place, right time and to the right people. Quality assessment of data in HIV prevention projects forms an integral part of monitoring and evaluating the performance of the projects. DQA is meant to motivate the client, staff and project stakeholders within and without the project as well as improve service delivery and this can only be achieved by disseminating the results of the DQA process (McKenzie, 2008). The process is seen as informative to different groups within the project and those not directly involved.

An assessment of data management and reporting systems through repeated assessments is used by the project team to identify gaps and monitor improvements made. In most countries, as far as DQA is concerned, data is usually collected using questionnaires to evaluate data collection and reporting. In Ethiopia, for example, the federal Ministry of Health introduced a new system to facilitate the DQA process which in turn facilitated improved measurement and standardization to ensure quality data and enable better decisions and better health outcomes as far as HIV/AIDS is concerned.

The uptake of HIV data from projects in any country is very high with the users of this data coming from a wide spectrum both in and out of the project scope. Most countries have tried to adopt a standardized data management, reporting and dissemination system in the functional levels ensuring that DQA be uniform across the board. DQA providing a basis for informing users how this data collected will be used is a fundamental role of the process. In the project itself, such data is used to know the key functional areas within the project and the areas the project managers need to put more effort in be it training of the data collection and analysis team, problems with the indicators used and many more.

2.4.1 Data Management System

DQA plays a role in providing information in all areas in any project as will be seen in the figure below. In a study done by Brackstone (1999) he notes that DQA looks at data quality as consisting of three dimensions which are the process, data and users. This empirical view sees to it that the DQA process ensures that the components of data quality specified at the beginning/inception of the project are captured. This in overall assists in facilitating assessment
and measurement and in providing a framework for stipulating guidelines and data quality improvement plans and ideas. According to the Data Quality Foundation, for the process to be informative, there are several aspects that have to be addressed in terms of the data in the various areas where it is collected. The first is the description of the data, which shows the connection between the facilities/point of collection and the completion of the source document that records the data. The documentation review which reviews completeness, availability and timeliness of all source documents for the reporting period. The other aspect is cross-checking of accuracy and reliability which relates to the verified recorded and reported numbers between data sources and finally spot checks are carried out which verify the actual data collected.

The data management system plays a very crucial role in the DQA process. It shows all the areas that the data from the DQA process is required. The information produced from the process is needed in all the sectors of the project for different purposes. The quality of reported data and use of the subsequent information is dependent on the underlying data management and reporting systems. Stronger systems are desired as they produce better quality data therefore for good quality data to be produced by and flow through a data management system, key functional components need to be in place at all levels of the system. The functional areas include the M&E areas, data collection, etc. The figure below shows areas in which DQA can be used to improve within a project. It shows the dimensions used to ensure quality of HIV data and the areas that can be improved after the process.

Figure 1: Components of a Data Management System
In a case study in Ethiopia, results from the data quality assessment enabled the researchers to understand qualitatively and quantitatively the operations as well as the strengths and weaknesses of functional areas that affect the overall quality of data in the health systems they were studying. Information from the assessment was used in the different functional areas and assisted in coming up with recommendations. Findings showed that standard data collection and reporting tools were in place. However, implementation was low and some of the factors were lading to this were inadequate provision of the required resources or inputs, including lack of trained focal persons, inadequate infrastructure. It is through such findings that led to the development of the Ethiopia Data Quality Assessment Framework (EDQAF).

2.4.2 Dissemination of Data Quality Assessment Information.

The informative process of DQA requires that the information needs of different stakeholders be assessed (Dhawan, 2000). A participatory approach is seen as the most suitable for such an assessment as it enables the relevant parties to clearly articulate their needs and judge how the particular content will be of use to them. In as far as DQA is concerned, Porter (1999) describes that the information emerging from the process can be packaged in three ways depending on the intended audience. The information can either be condensed to contain only concise details, it can be consolidated such that it contains the actual reports, findings and analysis or it can be repackaged which is normally done for the non-technical audience.

The DQA findings need to be disseminated by the project implementers and they will normally focus on utility standards which intend to ensure the assessment will serve all the information needs of all the intended users of the data. In the NASCOP projects, dissemination of the DQA reports is vital to the growth and sustainability of these projects. Most of the time reports will be distributed through the projects website, newsletters and bulletins, site discussions among the various parties, seminar meetings, emails and over the phone ensuring access to the information when needed.

The informative process of DQA allows for a focused data requirements and optimized design for data collection, use of clearly developed work plans for collecting data in the field, uniformly documented data collection, evaluation, and use, clearly developed analysis plans and sound, comprehensive quality assurance project plans. This in the long run enables the projects to run
smoothly and the overall findings are rapidly reviewed by regulators and other stakeholders, defensible results on which to base decisions, increased credibility with regulators and stakeholders and an overall better use of project resources. The reports require that the data be presented in simple form that is easy to interpret and use. They should also provide clear responses to the areas of data that were assessed as well as the findings and show clear steps for implementing improvements. Since these projects have many stakeholders who require different types of information the DQA process has a provision that facilitates the dissemination of findings to inform the users and allows them to make suggestions for improvement.

The preparation, dissemination and discussion of assessment reports and results to the various groups of audiences is seen as a final step in concluding the assessment including provisionally scheduling the next assessment. There are three main categories of reports prepared by the Quality Assessment Team for an assessment. Report for data producers and it contains a comprehensive report, an account of the assessment process, scores and comments for all quality dimensions covered by the assessment, an indication whether the data outputs are of sufficient quality to be official statistics, and detailed descriptions of the major quality problems and proposed quality improvements. Report for data users and comprises scores and comments for all quality dimensions covered by the assessment, and on whether the data are of sufficient quality to be made official. It also includes a summary of the major quality problems as they may limit use of the data, and an indication of any quality improvements that have been agreed. Report for the senior officials’ .This report is targeted for the senior officials and aims to give an impression of output and systems quality, whether the data is of sufficient quality to and to secure support for quality improvements. The report includes an overview of quality problems and quality improvement proposals and scores and comments for all quality dimensions covered by the assessment.

According to the WHO and UNAIDS research, the information after the DQA process should be disseminated in a procedural manner and they came up with steps suitable for nearly all projects and organizations. The steps included identifying target audiences and their information needs these include the partners/donors, project teams and the other stakeholders and collecting stakeholder information which involves knowing where and how the information will be sent. Once this is done, the project manager then determines the types of products that meet the audiences’ information needs since not all the information arrived at after the DQA process is
relevant to all; it is therefore the duty of the project manager to sieve and make sure relevant information is sent to relevant people. Afterwards the language requirements per products and audiences are identified, effective ways of dissemination are chosen and finally the results are measured and feedback is monitored.

In as much as information is important in such projects, most of them have had several impediments to this information dissemination agenda and they have been divided into three groups summarized by Yarvis 2001 as technical issues involving aspects such as the appropriateness of specified indicators, data collection methods, IT infrastructure, data presentation and personnel training, systematic & environmental issues include resources availed for the process, roles and responsibilities of staff and the overall organizational culture and behavioral issues which include aspects such as motivation, attitudes and values, confidence and a sense of ownership of the data being handled.
2.5 Standard Operating Procedures in Data Quality Assessment.

DQA has been described as a process with sequential steps and like any other process there are a defined set of rules that govern the process known as Standard Operating Procedures (SOP’s). According to the US Environmental Protection Agency, SOP’s are a set of written instructions that document a routine or repetitive activity. They facilitate conformance to technical and quality system requirements as well as enhancing and supporting data quality. Data quality has become an important issue not only because of its importance in promoting high standards of project deliverables but also because of its impact on government budgets for maintenance of health related projects (WHO, 2003).

DQA is a fairly new process; the initial steps in the original DQA had five steps from the review of the project objectives and a preliminary data review. It was then followed by selecting the statistical methods, verification of assumptions and then drawing conclusions (EPA Guide, 2006). According to Dan Solak (2000), there are three key stages that should be adopted by every project to enhance data quality; data capture which is the ability to aggregate all the relevant information collected, data processing which is the ability to process real-time inputs along with other relevant information to determine how best to respond to changing circumstances and data contribution, the ability to articulate your response to the users based on conclusions drawn from data processing.

2.5.1 Data Quality Assessment Standards and Procedures

Specifying standards and procedures for DQA facilitates assessment and measurement; it provides a framework for stipulating guidelines and data quality improvement plans provides a path for actions to improve data quality. (Loshin, 2001). In the DQA process there are three stipulated standards that are categorized as models (Veregin, 1998). The first is the Minimum Quality Standards which is a form of quality control where DQA is the responsibility of the data producer. It is based on compliance testing strategies to identify the data that meet the specified quality thresholds. An example is the National Map Accuracy Standards (NMAS) adopted by the US Geological Survey in 1946. There have been reservations to these standards since it is restrictive and lacks flexibility. The second category is the Metadata Standards which views error as inevitable and does not impose a minimum quality standard. It leaves the decision to the data...
users to determine quality of the data for themselves. An example is the Spatial Data Transfer Standard (SDTS). These standards are flexible but there is a one-way information flow that inhibits correction mistakes. The final category is the Market Standard which uses a two way information flow to obtain feedback from users on data quality problems. The feedback is processed and analyzed to identify significant problems and prioritize repairs.

A joint research done by the National Statistics Institutions in Canada, Netherlands, Korea, Eurostat and International Monetary Fund (IMF) came up with procedural guideline for conducting DQA which most countries around them have adopted for the process of DQA. Through peer review systems where the process used in one country are independently evaluated by experts from another country, they were able to establish similarities and the result was the adoption of the universal flowchart for DQA. By having such a universally used framework allows for uniformity and also for the entrenchment of specific standards in the area where the assessment is being carried out. The flowchart was found to assist in improving the assessment process and in turn the overall performance. From when the assessment is organized, the initial indicators to be assessed are selected and opinions sought on whether they are appropriate. In case they are not, they are revised and once in agreement the actual assessment is carried out and the findings deliberated upon and a final report prepared if there are no discrepancies. It also specifies that improvement be carried out once the report is delivered.

The flow chart on page 22 was developed by the Statistical organizations has served as the baseline or benchmark for which the operating procedures used in DQA have been established. The importance of properly established and managed quality control and quality assurance systems with their integral well-written SOPs and other quality documents for the achievement of project objectives greatly outweighs the time and resources used to develop them. These procedures enable success by assisting in the development of high-quality processes, procedures, systems, and people and may lead to benefits such as stakeholder satisfaction, timeliness of results, and alignment of the processes to the overall organization strategy and motivation of staff.
Figure 2: Flowchart for Universal Data Quality Assessment. *(Adapted from proceedings of Statistics Canada Symposium, 2001)*
According to the Australian Health Agency, before DQA was fully entrenched in the projects, reports indicated that there were some project managers who were involved in the manipulation of data and inflation of figures to appease the donors and increase donor funding to their projects. This is what led to having standard procedures defined and how they are followed. When key components in the DQA process are assessed individually looking at the specific indicators, it then becomes very easy to notice discrepancies and after a while of exposing the data to these checks and assessments, data quality is enhanced and project performance automatically improves since decisions made are more accurate, resources are allocated based on actual statistics and not assumptions.

In Africa, the major stakeholders in terms of funding for these HIV prevention projects are CDC, Global Fund and USAID. They have adopted an almost similar operating guideline for purposes of uniformity. According to the USAID document No. 18 in 2010, they provide that a DQA be conducted at least once every three years before the data is sent to the partner states mostly Washington. The process begins from identification of the DQA team, development of an overall approach and schedule, Identification of indicators to be included in the review and categorizing them. it is followed by holding working sessions to review indicators with the project team, holding sessions with implementation partners to review indicators, preparation of the DQA document and a follow up on the action plans.

The procedure for conducting the assessment is also well specified. A Data Quality Assessment (DQA) Checklist is provided as a recommended tool that an operating unit (OU) may use to complete its DQAs. There are several recommendations by USAID in conducting the assessment. The assessor should make sure that they understand the precise definition of the indicator by checking the Performance Indicator Reference Sheet any issues of ambiguity should be resolved before the DQA is conducted. The assessor should have a copy of the methodology for data collection in hand before assessing the indicator. For USAID Missions, this information should be in the PMP’s Performance Indicator Reference Sheets for each indicator. Each indicator should have a written description of how the data being assessed are supposed to be collected. Each implementing partner should have a copy of the method of data collection in their files and documented evidence that they are collecting the data according to the methodology. The assessor should record the names and titles of all individuals involved in the
assessment. The assessor should include a summary of significant limitations found. A plan of action, including timelines and responsibilities, for addressing the limitations should be made.

According to the Ethiopian Data Quality Framework, it uses a phased approach to data assessment where there are four assessment phases which are the Initiation and Preliminary Investigation phase which involves setting up the assessment schedule and working relationships with data producer and key users, conducting an initial review of metadata and other documentation and having preliminary discussions with the data producer, obtaining a sufficiently informed perception of data process quality to determine whether the next phase will be necessary and establishing the logistics of the assessment. The next phase is Systems Assessment which involves on site assessment at samples of zones and regions through which the data pass. It has two components which are the assessment of data collection, capture, processing and transmission procedures, and review of metadata and other documentation, at each level and verification of the data received and transmitted at each level. The third phase is the Overall National Assessment which involves assessment of output, institutional and process quality dimensions based on detailed discussions with production staff at national office, review of metadata and other documentation at national level and review of results of the second phase assessment if it was conducted. The final phase is the Reporting and Conclusion and it involves preparation of assessment results, comprising descriptions of major quality problems and potential quality improvements and quality summaries, distribution and discussion of these results with the data producer, senior managers having oversight of the data production process, and key data users and formally wrapping up the assessment process and provisionally scheduling the next assessment.

NASCOP has tried to customize it to suit the needs of its projects. It conducts DQA annually, depending on the need and objectives of the assessment or bi-annually to allow for any recommendations to be implemented. The DQA steps used by NASCOP as per the NASCOP Curriculum 2010 are determining the purpose of the audit, select levels and sites to be included, identify indicators, data sources, and reporting period, conduct site visits, review results and findings and the development a system-strengthening plan, including follow-up actions. There is a strong relation between the utilization of SOP’s in the DQA process and the performance of not only the HIV prevention projects but most projects that use this process.
2.6 Data Quality Assessment and Data Quality Improvement

Data quality improvement is an important element that is taken into careful consideration in the DQA process. In the steps of DQA mentioned in the previous section and the flowchart, the aspect of improvement has been incorporated in the process showing just how important it is. As the assessment is done, steps should be taken to improve the quality of data (Davis and LaCour, 2002). The most popular method used for quality improvement adopted by most countries is the Plan, Do, Check, Act (PDCA) method developed by Walter Shewart (Abdelhak, 1996).

**Figure 3: Plan-Do-Check-Act Cycle**

### 2.6.1 Plan-Do-Check-Act Cycle and Process Improvement

The PDCA cycle also known as PDSA cycle has been used to facilitate improvement of several different processes. The four phases in the cycle are continuous and each phase represents a set of activities that facilitate improvement and has been applied in the improvement of the DQA process. The PLAN phase of the cycle consists of looking at the improvement opportunity which is usually specified at the end of the assessment. The current process used is discussed as the project team tries to determine which aspects do not work and is measured and if there is no specific or special cause of the problem the process is maintained but if a discrepancy is discovered the change is the planned and specified. In the DO phase the process is undertaken in
the new way and monitored. In the CHECK phase, the methods of data collection and the data collected are verified and the impact of the change, if any, is studied. In the ACT phase the documentation of the process is done and the improvements are incorporated in the process.

The assessment process facilitates the process of data quality improvement through various ways. According to a WHO publication in 2003, the development of performance indicators, policies on data, coming up with a quality review team within the project all of which are done in DQA contribute to the improvement process. Larisse, et al, (2005) indicated that the use of technology in the assessment process raises the level of having quality data for purposes of improvement; DQA requires a stronger personal involvement in the process, incentives, enforcement of policies, training and other injections to ensure continuous improvement. In the DQA process, NASCOP has realized that the improvement aspect of data is essential. The data being assessed during the process is data that leads to identification of trends, understanding causes and alternative solutions as well as tracking improvement. Through the use of the PDCA cycle, this has been almost realized although more still needs to be done. An improvement strategy is set up that facilitates the continuous improvement agenda through fostering a data quality culture, strengthening data quality infrastructure and capacity, cultivating the data supply chain, enhancing external data quality collaboration, initiating fast-track priority projects and promote communication and provide consultation.

Improvement is always gradual and undertaken in small steps to check whether the solutions being implemented are actually making a difference in the process. It begins with the brainstorming of ideas and theories that may cause a change and when a viable idea is found it is tested in a small scale. A follow-up test is then done to check for consistency of results and if it still works the idea is implemented in a wide scale and finally the change is finally incorporated in the process or system as illustrated in the diagram below.
In an article by Werner Engelen, the process of improvement is continuous and he arrives at several often overlooked procedures that lead to improvement especially in a project setting. Once a DQA environment is established within a project, it is the duty of the project managers to assess data definitions and ensure quality data is collected. The issues arising during the assessment should be identified and the impact they have on the overall project analyzed. Once these are sought, it is then imperative to analyze the causes of the issues, implement remedies and monitor to see if there are any changes. In improvement, the first aspect involved is to define the data quality problem by establishing the scope of the project, its objectives and the criteria to be used to check whether there is conformance to the data quality standards specified. It is followed by measuring the conformance to data quality standards and recording exceptions against the established data standards. The next step is to analyze conformance and prioritize conformance issues so as to give recommendations for improving data quality and finally improving data quality by implementing recommendations.

### 2.6.2 Improvement Methods in Data Quality Assessment

According to a research by Bunney, 1999, there are various ways that improvement is facilitated through the DQA process. Ensuring routine data quality checks as part of on-going supervision: For example, routine data quality checks can be included in already planned supervision visits at
the service delivery sites. Use of initial and follow-up assessments of data management and reporting systems: For example, repeated assessments (e.g., biannually or annually) of a system’s ability to collect and report quality data at all levels can be used to identify gaps and monitor necessary improvements. Strengthening program staff’s capacity in data management and reporting: For example, M&E staff can be trained on the DQA and be sensitized to the need to strengthen the key functional areas linked to data management and reporting in order to produce quality data. Preparation for a formal data quality audit: The DQA tool can help identify data quality issues and areas of weakness in the data management and reporting system that would need to be strengthened to increase readiness for a formal data quality audit. External assessment by partners of the quality of data: Such use of the DQA for external assessments could be more frequent, more streamlined and less resource intensive than comprehensive data quality audits that use the DQA version for auditing.

According to Redman 1992, a major activity in improving data quality is to execute the recommendations and monitor the implementation. Improvement work tends to fall into four categories. Process Improvement is the first and it focuses on improving the functional processes that are used to create, manage and use data e.g. introducing functional processes that encourage centralized data entry. The next is system improvement which has aspects such as software, hardware, and telecommunication changes that can aid in improving data quality. They may include providing better end-user manuals, operation and maintenance manuals, and additional user training. Policy and Procedure Improvement is the next category and involves resolving conflicts in existing policies and procedures and developing appropriate guidance that will institutionalize the behaviors that promote good data quality and a major example already discussed earlier is the development of Standard Operating Procedures (SOP) that document the data quality rules and sets filters that are used to measure data quality. In addition, the performance of periodic data quality checks may be performed as part of the SOP to increase data quality. The final category is data design improvement which aims in improving the overall data design and using the specified organization data standards which ensure the uniform representation of data across all the projects and supports improvements in data correlation.
2.7 Conceptual Framework

The figure below shows the relationship between the independent and dependent variables and gives the appropriate indicators for each.

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Moderating Variable</th>
<th>Dependant Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Collection &amp; Informative Process of DQA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Availability of data collection tools</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Relevance of data collection tools</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• No. of meetings</td>
<td>NGO Policy</td>
<td></td>
</tr>
<tr>
<td>• No. of site discussions</td>
<td>Health policies</td>
<td></td>
</tr>
<tr>
<td>• Availability of a website</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Regular dissemination of brochures</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Standard operating procedures in DQA</strong></td>
<td></td>
<td>Performance of HIV Projects</td>
</tr>
<tr>
<td>• Availability of the SOP’s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Effectiveness of procedures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Relevance of procedures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Implementation levels</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Data Quality Improvement in DQA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• No. of trainings in DQA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• No. of new services introduced</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• No. of data quality issues in DQA report</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Intervening Variable</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Figure 5: Conceptual Framework</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.8 Conclusion
The introduction of DQA has led to the achievement of several milestones towards achieving the data quality objective. The self assessments provide an overall picture of how the project phases conform to the specified standards. The DQA process has been so revolutionary that ISO standards as regards to data quality and the processes to ensure this have been developed. DQA is a very expensive process that is not necessarily a must in any project which makes them opt not to carry it out but as we go on, the effects of poor quality data have been felt in some industries and this helps to create awareness to others to embrace the process not as a separate process in the project but as an iterative and continuous process throughout the project cycle. DQA acts as a process of eliminating the jargon in the data collected and analyzed and presenting it in a form that is easily understood to the people not directly involved in the project. The findings and DQA reports show how all the data that was collected will be used and the type of decisions the information will facilitate in making. The presence of these standards and procedures eliminate mistakes and aspects such as redundancy in the process ensuring that quality is achieved and maintained. They also provide a final check for the entire project and its data to ensure all the steps required were followed and the outcome is of quality and acceptable standards and that it leads to the achievement of the overall objectives of the project.
CHAPTER THREE
RESEARCH METHODOLOGY

3.1 Introduction

In this chapter the researcher discussed the research design to be used and why it was the most appropriate for the study, the targeted population as well as justified the sampling procedure and the sample size used in the study. It also looked and justified the use of the data collection methods, reliability and validity of the tools to be used for data collection as well as the procedures used for analysis.

3.2 Research Design.

This study employed a descriptive survey research design. This was seen as the most appropriate research design as it helped the researcher describe the behavior of the population that was under study and allowed for generation of both qualitative and quantitative data. It was chosen due to the fact that it allowed for the information on the projects to be collected and look at how the aspects of informative process, standard operating procedures and quality improvement affect performance. This research design enabled the researcher explain the conditions that were happening and allowed the researcher to ask questions to the respondents and a chance to summarize the findings by observing and describing the population under study without any interference.

3.3 Target Population

The population for this study was staff in NASCOP prevention projects that deal with the Most at Risk Populations (MARPS), HIV testing and counselling and comprehensive care in Nyeri County. The population was drawn from interventions undertaken in Nyeri Town Center, Mathira, Othaya, Kieni and Mukurweini. In the five areas the study targeted the District Aids Co-coordinators, data clerks, data analysts and data collection staff. It also included the Data Management Assistant, and the Provincial Aids Co-coordinator who are from the Ministry of Public Health and NASCOP respectively.
3.4 Sampling Procedure and Sample Size

3.4.1 Sampling Procedure

The researcher used both probabilistic and non-probabilistic methods of sampling. The cluster sampling technique was used. The project areas in Nyeri County were viewed as clusters. Once the clusters were identified, a simple random sampling technique was used to get the respondents in these clusters. Purposive sampling was used for specific categories of respondents and it was up to the researcher to select which respondent is able to give the most accurate information. This was based on looking at which projects have been most successful.

3.4.2 Sample Size

According to the Pennsylvania State University article on Program evaluation (2007), there is no set percentage that is accurate for every population when selecting a sample size, what matters is the actual number or size of the sample, not the percentage of the population. The decision on sample size for this study was based on the objectives of the research, the expected precision of the results and the estimated response rate of the participants seeing as the duration of the research is limiting. The table below shows the distribution of the population and the estimated sample size to be used:

Table 3.1: Population Size

<table>
<thead>
<tr>
<th>Category of Respondents</th>
<th>Total population in all the project areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data clerks</td>
<td>70</td>
</tr>
<tr>
<td>Data Analysts &amp; project data team</td>
<td>56</td>
</tr>
<tr>
<td>Project Managers</td>
<td>5</td>
</tr>
<tr>
<td>DMA</td>
<td>8</td>
</tr>
<tr>
<td>DHRIO</td>
<td>3</td>
</tr>
<tr>
<td>PASCO</td>
<td>1</td>
</tr>
<tr>
<td>TOTAL</td>
<td><strong>143</strong></td>
</tr>
</tbody>
</table>

The formulae for calculating sample size will be used to determine sample size.
\[ n = \frac{N}{1 + N(e)^2} \]

Where \( n \) = the desired sample size

\( N \) = the population size. In this case 143

\( e \) = level of precision. In this case 0.07

\[ n = \frac{143}{1 + 143(0.07)^2} = 143 \]

\[ 1 + 143(0.07)^2 = 1.7007 \]

\[ = 84 \]

The sample size therefore consisted of 84 respondents.

### 3.5 Research Instruments

The study employed the use of questionnaires and an interview guide for the purposes of data collection. The reason for the use of questionnaires to collect data was that the researcher was able to collect the information in a standardized way and there was uniformity of results due to the respondents receiving the same questions yielding more comparable data which made it easier to analyze and also given the time sensitiveness of the research, they were easy to administer. The interview guide was used as it enabled the collection of more in depth information from the respondents to gain an in-depth understanding to how they usually conduct the process. The use of interview was chosen since it will give the researcher a chance to clarify questions not understood and also observe the non-verbal responses.

Data was collected from the data clerks because they are directly involved with data collection and use of the data collection tools therefore they are able to articulate the challenges they face better. The data analysts and the project managers were interviewed as they are involved in the DQA process. The research instruments were both open and closed ended to allow for maximum responses to the questions. The researcher administered these instruments personally to the
respondents and where this was not possible help was sought from research assistants. The instruments were simple to understand and instructions on how the questionnaires would be filled were given. The respondents were asked to ask the researcher in case they encountered any problems during this process.

Apart from these, the researcher also employed some observation during the process as well as use of secondary data from previous reports produced after the process was done and this enabled a comparison to be made about the process then and now.

3.6 Pilot Study

Once the instruments were developed, they were administered to a sample with similar characteristics as the one that will be used for the study. The researcher chose Kiambu county and more specifically Thika District due to the fact that the projects undertaken are similar to those in Nyeri County. This helped the researcher to ascertain whether the instruments were able to deliver the desired results. The purpose of the pilot study was to pre-test the research instruments that were developed by the researcher. This served to give an advance warning about where the main research project could fail, where research protocols may not be followed, or whether proposed methods or instruments are inappropriate or too complicated. It also served to check the validity and reliability of the developed instruments. The researcher managed to get 20 respondents to participate in the pilot test.

The research instruments were administered to this pilot group and the responses were varied. Form this pilot, the researcher was able to re-evaluate some of the questions asked as some of the respondents found them ambiguous and also incorporated some of their suggestions in the final research instruments.

3.6.1 Validity of Research Instruments

Validity determines whether the research truly measures that which it was intended to measure or how truthful the research results are (Joppe, 2000). The validity of the instruments developed was determined through consultative discussion between the supervisor and the researcher. Through this, the researcher was able to see whether the developed instruments would measure the concept the researcher intended.
3.6.2 Reliability of Research Instruments

The extent to which results are consistent over time and an accurate representation of the total population under study (Joppe, 2000). The instruments should ensure that there is some form of consistency to ensure that the research is of quality.

Reliability of the instruments was tested through the test-retest method to check whether the participants’ responses would change over time. Once the instruments were developed, the researcher administered the instruments to a population similar to the one to be used in the study. The data was collected, analyzed and the findings recorded and after two weeks administered the same instruments to the same participants to check whether the same results would be achieved as would be expected. A correlation coefficient was calculated to determine how closely the participants’ responses on the second time matched those given the first time. A correlation coefficient between 0 and 1 was expected. If the co-efficient was below 0.5, these instruments would be seen as unreliable and the researcher would back to develop new instruments. If the co-efficient is above 0.5, the instruments would be ruled as reliable and the actual data collection could start. The correlation coefficient obtained from the reliability test was 0.636 which is a positive correlation. This meant that the responses given the first time the tools were administered were relatively similar to those received the second time.

3.7 Data Collection Procedure

After the instruments had been developed and the pilot test done, the researcher then administered the questionnaires first as they required more time due to the travelling involved. The researcher went to most of the areas and administered the instruments while in other parts the research assistant was used. Once all the questionnaires had been administered to the respondents, they were advised to place them in the office for ease of collection. The researcher then planned on a day to conduct the interviews with the respective respondents. The interviews were used to get in depth information that could not be known through the questionnaires and were all carried out in two days. Once all the data collection tools were administered and returned to the researcher, they were stored entered and stored in the computer.
3.8 Data Analysis

The process of data analysis began with data cleaning which was done so as to minimize the number of errors likely to be encountered. Incomplete questionnaires and any other inconsistencies were eliminated. The researcher then organized the data by coding the responses by assigning values to the responses. The data was then fed to the computer and the SPSS software program which was used for analysis. The necessary calculations were performed and the findings presented in tabular form. The process of interpreting the data to a meaningful form followed. Once the data had been interpreted, the researcher discussed limitations and drew conclusions on the research from the data.

For the qualitative data, the data was analyzed through reading and reviewing the interviews. A unit of analysis was then determined by looking for keywords and phrases and afterwards coding the data. Once this was done the data was interpreted and reported and meaningful conclusions drawn which was how the content analysis was performed. Through this content analysis the researcher was able to capture all the responses and collate them into one meaningful format.

3.9 Ethical Considerations

The researcher while undertaking this study was guided by the three principles of ethics which are respect, beneficence and justice. The researcher ensured that a letter of authority to collect data was received from the National Commission for Science and Technology (NACOSTI) and the University. Other aspects of ethics that were considered were to ensure that the principle of voluntary participation was followed by explaining the purpose of the research, expected duration and procedures as well as informing the participants' of their rights to decline to participate and to withdraw from the research once it has started. A letter of introduction from the employer was also used to confirm to the participants who the researcher was. The researcher ensured that confidentiality of the responses was maintained by not having them indicate their names on the questionnaires and issuing the questionnaires to individuals to avoid collusion. Also, having a password on the computer where data was being analyzed to prevent findings from being tampered with.
<table>
<thead>
<tr>
<th>Objectives</th>
<th>Type of Variable</th>
<th>Indicators</th>
<th>Measurement Scales</th>
<th>Research Instruments</th>
<th>Method of Data Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Role of DQA in the performance of HIV prevention projects</td>
<td>Dependent</td>
<td>Performance of HIV projects</td>
<td>Ratio</td>
<td>Interview, questionnaires</td>
<td>Descriptive statistics</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Number of completed projects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Achievement of objectives</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Cost effectiveness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Appreciation by recipient population</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Efficiency of services</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>To investigate how SOP’s in DQA influence performance of HIV projects in NASCOP.</td>
<td>Independent</td>
<td>SOP’s in DQA</td>
<td>Ratio</td>
<td>Interview</td>
<td>Descriptive statistics</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Availability of the SOP’s.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Accuracy and completeness of data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Effectiveness of Procedures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Relevant procedures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Proper implementation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>To determine how the informative process in DQA influences HIV projects in NASCOP.</td>
<td>Independent</td>
<td>Informative Process of DQA</td>
<td>Nominal</td>
<td>Interview and questionnaires</td>
<td>Descriptive statistics</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• No. of meetings</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• No. of site discussions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Availability of a website</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Regular dissemination of brochures</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.11 Dissemination Plan

The dissemination plan explains the purpose of the dissemination, the message that will be disseminated and to whom the dissemination will target, the method of dissemination and the specific time the dissemination will take place. For this research study, the researcher chose to disseminate the findings by publishing the findings in a selected journal and workshops. The target for this dissemination will be the project managers and the data management teams and is expected to take place two months after the completion of the research study. A brief report will be distributed during the workshops and is expected to start two months after the completion of the research. It is estimated to cost up to Ksh. 50,000 and will be evaluated through the number of citations the journal receives and the feedback from the workshop audience.

3.12 Summary

This section of the study looked at the aspect of data collection and analysis. It specified the methodology to be used for data collection, the sampling procedure and sample size. It also looked at the research instruments to be and how their reliability and validity will be ensured. The collection and analysis of data were also discussed in this chapter as well as the ethical issues that were under consideration during the data collection exercise.
CHAPTER FOUR

DATA ANALYSIS, PRESENTATION AND INTERPRETATION

4.1 Introduction

This chapter focused on data analysis, interpretation and presentation and was guided by the objectives and the research questions. The objectives of the study were to assess how data collection and the informative process in DQA influences performance of HIV projects, to investigate how Standard Operating Procedures in DQA influence performance of HIV projects and to establish how data quality improvement in DQA influences performance of HIV projects in NASCOP.

4.2 Response Rate

The response rate of the 2 categories of respondents is presented in Table 4.1

**Table 3.1: Response Rate**

<table>
<thead>
<tr>
<th>Category</th>
<th>Sample Size</th>
<th>Response</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project managers and data analysts</td>
<td>5</td>
<td>5</td>
<td>100</td>
</tr>
<tr>
<td>Data clerks, DASCOs and HRIOs</td>
<td>79</td>
<td>73</td>
<td>92.4</td>
</tr>
</tbody>
</table>

Table 4.1 illustrates the response rate of the respondents who were sampled and interviewed in the study. The study targeted 79 data clerks, DASCOs and HRIOs and 5 project managers and data analysts. The response was 92.4% for data clerks, DASCOs and HRIOs and 100% for Project managers and data analysts sampled completely filled in and returned the questionnaire. This is attributed to the fact that the researcher employed 2 research assistants to personally administer and guide the respondents in answering the questions. The return rate which is over 50% meant that the findings were more accurate.
4.3 Background Information of the Data clerks, DASCOS and HRIOs

In this section the researcher sought to establish the gender of the respondents and looked at their gender, age, education level and position.

4.3.1 Gender of the Respondents

In this section the researcher asked the respondents to indicate their gender. The responses are shown in Table 4.2 and Table 4.3.

Table 4.2: Gender of Data Clerks, DASCOs and HRIOs

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>35</td>
<td>47.9</td>
</tr>
<tr>
<td>Female</td>
<td>38</td>
<td>52.1</td>
</tr>
<tr>
<td>Total</td>
<td>73</td>
<td>100.0</td>
</tr>
</tbody>
</table>

From the Table 4.2, 47.9% of the Data clerks, DASCOS and HRIOs were males while the rest were females. This implied there is a small gender disparity between the responses of males and females and that the project considered both men and women equally capable of performing the duties assigned to them.

Table 4.3: Gender of Project Managers and Data Analysts

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>5</td>
<td>100</td>
</tr>
<tr>
<td>Female</td>
<td>0</td>
<td>.0</td>
</tr>
<tr>
<td>Total</td>
<td>5</td>
<td>100.0</td>
</tr>
</tbody>
</table>
From the Table 4.3 all the project managers and data analysts were males and this showed a gender bias in as far as project heads are concerned. The projects should consider hiring more women for gender balance.

4.3.2 Age of the Respondents

The researcher asked the respondents to indicate their age. The responses are shown in Table 4.4 and Table 4.5.

Table 4.4: Age of Data Clerks, DASCOs and HRIOs

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>18 years to 24 years</td>
<td>4</td>
<td>5.5</td>
</tr>
<tr>
<td>25 years to 30 years</td>
<td>29</td>
<td>39.7</td>
</tr>
<tr>
<td>31 years to 35 years</td>
<td>18</td>
<td>24.7</td>
</tr>
<tr>
<td>36 years to 40 years</td>
<td>8</td>
<td>11.0</td>
</tr>
<tr>
<td>46 years to 50 years</td>
<td>5</td>
<td>6.8</td>
</tr>
<tr>
<td>Over 50 years</td>
<td>9</td>
<td>12.3</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>73</strong></td>
<td><strong>100.0</strong></td>
</tr>
</tbody>
</table>

From the Table 4.4, 69.9% of the data clerks, DASCOs and HRIOs were aged 35 years and below while 30.1% of the data clerks, DASCOs and HRIOs were aged 36 years and above. The majority of the respondents between 25 years to 40 years showed that the projects have a mature workforce capable of undertaking the duties assigned to them.
Table 4.5: Age of the Project Managers and Data Analysts

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>18 years to 29 years</td>
<td>0</td>
<td>.0</td>
</tr>
<tr>
<td>30 years to 39 years</td>
<td>4</td>
<td>80.0</td>
</tr>
<tr>
<td>40 years to 49 years</td>
<td>1</td>
<td>20.0</td>
</tr>
<tr>
<td>50 years to 59 years</td>
<td>0</td>
<td>.0</td>
</tr>
<tr>
<td>60 and above years</td>
<td>0</td>
<td>.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>5</strong></td>
<td><strong>100.0</strong></td>
</tr>
</tbody>
</table>

From the Table 4.5, 80.0% of the project managers and data analysts were aged between 30 years and 39 years and 20.0% of the project managers and data analysts were aged between 40 years and 49 years. This shows that all the respondents were mature and had the ability to grasp the concept of the study and are also very capable of undertaking their duties diligently.

4.3.3 Level of Education of Respondents

In this part the researcher asked the respondents to indicate their level of education. The responses are shown in Table 4.6 and Table 4.7.
### Table 4.6: Level of Education of Data clerks, DASCOs and HRIOs

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td>0</td>
<td>.0</td>
</tr>
<tr>
<td>Secondary</td>
<td>3</td>
<td>4.1</td>
</tr>
<tr>
<td>Certificate</td>
<td>2</td>
<td>2.7</td>
</tr>
<tr>
<td>Diploma</td>
<td>41</td>
<td>56.2</td>
</tr>
<tr>
<td>Degree and above</td>
<td>27</td>
<td>37.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>73</strong></td>
<td><strong>100.0</strong></td>
</tr>
</tbody>
</table>

From the Table 4.6, 95.9% of the data clerks, DASCOs and HRIOs had post-secondary education while on 4.1% of the data clerks, DASCOs and HRIOs had secondary education. The respondents who had a diploma and above at 93.2% showed that the respondents have the necessary skills and are competent in performing the work they do.

### Table 4.7: Level of Education of Project Managers and Data Analysts

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td>0</td>
<td>.0</td>
</tr>
<tr>
<td>Secondary</td>
<td>0</td>
<td>.0</td>
</tr>
<tr>
<td>Certificate</td>
<td>0</td>
<td>.0</td>
</tr>
<tr>
<td>Diploma</td>
<td>0</td>
<td>.0</td>
</tr>
<tr>
<td>Degree and above</td>
<td>5</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>5</strong></td>
<td><strong>100.0</strong></td>
</tr>
</tbody>
</table>
From the table 4.7, all the project managers and data analysts had a degree and above. This shows that the respondents had some educational background in their respective field that was useful in their work experience and indicated some level of competency and skill which in turn is reflected in the performance of the projects.

4.3.3 Positions Held

Here, the researcher asked the respondents to indicate the positions held. The responses are shown in Table 4.8 and Table 4.9.

Table 4.4: Positions Held by Data clerks, District Aids and STI Co-coordinators and Health Records Information Officers

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data clerks</td>
<td>31</td>
<td>42.5</td>
</tr>
<tr>
<td>DASCO</td>
<td>23</td>
<td>31.5</td>
</tr>
<tr>
<td>HRIO</td>
<td>19</td>
<td>26.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>73</strong></td>
<td><strong>100.0</strong></td>
</tr>
</tbody>
</table>

From the Table 4.8, 42.5% of the respondents indicated they worked as data clerks, 31.5% of the respondents were DASCOs and 26.0% of the respondents were HRIOs. This showed that there was a distribution of labor among the data collection staff which in turn made the data collected more substantive.

Table 4.5: Positions Held by Project managers and data analysts (n=5)

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project manager</td>
<td>3</td>
<td>60.0</td>
</tr>
<tr>
<td>Data analyst</td>
<td>2</td>
<td>40.0</td>
</tr>
</tbody>
</table>
From the table 4.9, 60.0% of the respondents were project managers while 40.0% of the respondents were data analysts. From the figures, there seemed to be a deficiency in the staff that performs DQA where the data analysts seemed to have a heavier workload that the rest based on their numbers.

4.4 Data Collection and the Informative Process

From this, the researcher sought to address the first objective that assessed how data collection and the informative process in DQA influence performance of HIV projects in NASCOP. The researcher asked the respondents questions on how often they collected project related data, if they had a project website and to rate the aspects related to data collection and the informative process.

4.4.1 Frequency of collecting Project Related Data

Here, the researcher asked the respondents to indicate the frequency of collecting project related data. Their responses are shown in Table 4.10

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily</td>
<td>14</td>
<td>19.2</td>
</tr>
<tr>
<td>Weekly</td>
<td>28</td>
<td>38.4</td>
</tr>
<tr>
<td>Fortnightly</td>
<td>2</td>
<td>2.7</td>
</tr>
<tr>
<td>Monthly</td>
<td>29</td>
<td>39.7</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>73</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

From the Table 4.10, 39.7% of data clerks, DASCOs and HRIOs indicated that the collected project related data monthly, 38.4% of the data clerks, DASCOs and HRIOs collected the
projected related data weekly, 19.2% of the data clerks, DASCOs and HRIOs collected the
projected related data daily and 2.7% of the data clerks, DASCOs and HRIOs collected the
projected related data fortnightly. From the findings above, the project data is collected
throughout with the most being weekly and monthly at 78.1%. This increased the quality of the
data that is collected which in turn is reflected in improved performance of the project.

4.4.2 Project Website

In this section the researcher asked the respondents to indicate if they had a project website.
Their responses are shown in Table 4.11

Table 4.7: Project Website

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>25</td>
<td>34.2</td>
</tr>
<tr>
<td>No</td>
<td>48</td>
<td>65.8</td>
</tr>
<tr>
<td>I don’t know</td>
<td>0</td>
<td>.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>73</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

From the Table 4.11, 65.8% of the data clerks, DASCOs and HRIOs indicated they did not have
a project website while 34.2% of the data clerks, DASCOs and HRIOs indicated they had a
project website. The absence of a project website was seen as an indicator that affected
performance since the dissemination of DQA findings was hindered and this affected
performance.

4.4.3 Rating of Aspects Related to Data Collection and the Informative Process

In this section the researcher asked the respondents to rate the following aspects related to data
collection and the informative process using a likert scale of 1 to 4 with 1 being strongly agree, 2
being agree, 3 being disagree and 4 being strongly disagree. Their responses are shown in Table
4.12.
Table 4.8: Rating of Aspects Related to Data Collection and the Informative Process

<table>
<thead>
<tr>
<th>Category</th>
<th>S.A</th>
<th>A</th>
<th>D</th>
<th>S.D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>%</td>
<td>F</td>
<td>%</td>
</tr>
<tr>
<td>The project provides adequate data collection tools</td>
<td>23</td>
<td>31.5%</td>
<td>48</td>
<td>65.8%</td>
</tr>
<tr>
<td>The data collection tools clearly capture the relevant information with regards to the project.</td>
<td>25</td>
<td>34.2%</td>
<td>45</td>
<td>61.6%</td>
</tr>
<tr>
<td>The data collected adds value to the project as far as achieving project objectives is concerned.</td>
<td>32</td>
<td>43.8%</td>
<td>35</td>
<td>47.9%</td>
</tr>
<tr>
<td>The project management team adequately facilitates the data collection exercise?</td>
<td>6</td>
<td>8.2%</td>
<td>48</td>
<td>65.8%</td>
</tr>
<tr>
<td>The project website helps in improving work quality in terms of disseminating project related information.</td>
<td>24</td>
<td>32.9%</td>
<td>24</td>
<td>32.9%</td>
</tr>
<tr>
<td>Site discussions help improve quality of data collected and help in disseminating information.</td>
<td>29</td>
<td>39.7%</td>
<td>44</td>
<td>60.3%</td>
</tr>
<tr>
<td>Dissemination of project information is seen as a way of measuring project success.</td>
<td>41</td>
<td>56.2%</td>
<td>29</td>
<td>39.7%</td>
</tr>
</tbody>
</table>

S.A – Strongly Agree; A – Agree; D – Disagree; S.D – Strongly Agree; F – Frequency, % - Percentage

From the Table 4.12, 97.3% of the data clerks, DASCOs and HRIOs agreed with the statement that the project provides adequate data collection tools which leads to adequate amounts of data being collected and this in turn reflects on the performance of the project. 95.8% of the data
clerks, DASCOs and HRIOs agreed with the statement that the data collection tools clearly captures the relevant information with regard to the project but in as much as this was a large number, the researcher felt that the team that develops these tools should get more views on what the data clerks feel is relevant to the projects. 93.5% of the data clerks, DASCOs and HRIOs agreed with the statement that the data collected added value to the project as far as achieving project objectives is concerned, 74.0% of the data clerks, DASCOs and HRIOs agreed with the statement that the project management team adequately facilitates the data collection exercise and a large number felt that not much was being done to facilitate the dissemination. The motivation through facilitation of work done was seen as a factor that affected the project performance since the people behind the data need to feel motivated. 65.8% of the data clerks, DASCOs and HRIOs agreed with the statement that the project website helps in improving work quality in terms of disseminating project related information, all the data clerks, DASCOs and HRIOs agreed with the statement that site discussions helped improve quality of data collected and help in disseminating information which was a factor that influenced performance through the dissemination of findings. 95.9% of the data clerks, DASCOs and HRIOs agreed with the statement that dissemination of project information is seen as a way of measuring project success.

4.4.4 Ways Used to Ensure the Findings are known

The researcher asked the respondents to indicate the ways used within the project environment to ensure findings are known to all the stakeholders. In their response they said through monthly meetings, posters, quarterly stakeholders meetings, monthly service providers meetings, periodical meetings, through barazas, workshops and seminars, holding multi-disciplinary team meetings and sharing during CMEs and findings with other departments and stakeholders, publications, focus discussions, media, brochures, briefs, holding shareholders forum, publicizing the findings, involving the government to transmit the data using the administration, on-line social media, bulletins, organizing and attending project workshops, using print media, MDT’s, RRI discussions, emails, data review meetings, feedback meetings, on the job training and continuous medical education. With these different ways of disseminating information from DA, the researcher felt that this was a major tool the projects used to gauge performance. It was
clear that the projects teams understood the importance of disseminating the information they collected and that determined by the DQA process.

**4.4.5 Ways Project Management Teams can help in ensuring the Data Collection and Information Dissemination is more effective**

The researcher asked the respondents to indicate the ways they thought the project management team can help in ensuring the data collected and information dissemination is more effective. In their response they indicated that giving regular feedbacks for data collected, quarterly meetings and data quarterly review meetings, facilitate movement to access facilities where data is generated and the website is made more effective, provide a more regular way of dissemination information and have uniform tools, invest in ICT and recent technology, holding meetings at grass root levels for brainstorming sessions and gathering feedback from these meetings, training data collectors, have a regular way of disseminating information and facilitate the dissemination exercise, enlighten respondents of the importance of the said project, disseminating information through websites and emails, setting up website and vigorously advertising it to the community, providing adequate data collection tools, mentorship, tailor made data collection tools, increasing the frequency of collecting the data, helping the staff be conversant with data collection tools and giving collect data, support DASCOs more on supervision to the sites, timely dispatch of emails and all communication, timely allocation of resources, computerizing all data point to ensure timely collection of data and to improve accuracy, supporting the IT team with superior technology and capacity building service providers on EMR, workshops, facilitate DASCOS and HRIOs for movement, mentorship programs, availing data clerks, training all HRIOs on HMIS tools, weekly data analysis and monthly management meetings with data reports and reviews. Form this the researcher deduced that the project team and staff felt that technology would greatly help in making their work more effective and efficient which in turn would play out n an improved performance.

**4.4.6 Average Duration of the Projects Undertaken**

In this section the respondents were asked to indicate the average duration of the project they undertook. Their responses are shown in Table 4.13
Table 4.9: Average Duration of the Projects Undertaken

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 month to 4 months</td>
<td>0</td>
<td>.0</td>
</tr>
<tr>
<td>5 months to 8 months</td>
<td>0</td>
<td>.0</td>
</tr>
<tr>
<td>9 months to 12 months</td>
<td>0</td>
<td>.0</td>
</tr>
<tr>
<td>Over 1 year</td>
<td>5</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>5</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

From the Table 4.13 all the project managers and data analysts indicated the average duration of the projects undertaken was over 1 year. This showed that they are long term projects that were data intensive and therefore there was a need for the DQA to ensure that quality was not compromised in the duration of the projects. From these findings also, it was clear that since the projects are long term, the team tasked is competent and has the skills due to the repetitive nature of the tasks.

4.4.7 Critical Issues from the Data Collected

At this point the researcher asked the respondents to indicate the most critical issues from the data they collected that contribute to data quality in the projects. In their response, the project managers and analysts indicated that they valued integrity since HIV is still a touchy subject in most quarters. They also ensure reliability since the procedures they used are consistent. Knowledge and skills of data collectors, data review in use, data analysis and information dissemination, data obstruction, data cleaning methods, technical reviews, data accuracy and timeliness in reporting were all noted as critical issues in data collection that contribute to quality data. From the responses it was clear that there was presence of an assessment plan due to the indicators they looked at and their suitability for the type of projects that are undertaken.
4.4.8 Important components of quality Data that are assessed in the Projects

The researcher asked the respondents to indicate the most important components of quality data that were assessed in the projects. In their response the project managers and data analysts indicated that they assess the operational heterogeneity of the data, how accurate the data is and other aspects such as timeliness of the data. They also assessed the indicators to ensure that they are met and the accuracy, analysis, reviews, usefulness in decision making, quality, comprehensiveness and precision of the data were all important components.

4.4.9 Management of Expectations of Data Quality from Various Audiences

In this section the researcher asked the respondents to indicate how they managed the expectations of data quality from the various audiences. In their responses the project managers and data analyst indicated that they ensured that they had continuous flow of information at all times to keep them informed and also tried to have all the data reach the targeted audience once analyzed, others held several stakeholders meetings to disseminate the information and forums with different audiences to try and find out what aspects of data are most important to them. They also ensured data availability, dissemination of information to the targeted audience, had reviews from various groups and had data consumer forums with different audience.

4.4.10 Most Common Data Problems Encountered

Here the researcher asked the respondents to indicate from their experience what the most common data quality problems they encountered were. In their response the respondents indicated that the most common data quality problems they had encountered were abstraction, validation, staff turnover, incompleteness, inaccurate data, data collected, inconsistence and incomplete data as well as wrong analysis of data, incomplete tools with information missing and transcription errors during data collection.

4.4.11 Frequency of Disseminating Information from the Assessment

The respondents were asked to indicate how often they disseminated information from the assessment. Their response is shown in Table 4.14.
Table 4.10: Frequency of Disseminating Information from the Assessment

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Once in the duration of the project</td>
<td>0</td>
<td>.0</td>
</tr>
<tr>
<td>Twice in the duration of the project</td>
<td>3</td>
<td>60.0</td>
</tr>
<tr>
<td>Others</td>
<td>2</td>
<td>40.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>5</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

From the Table 4.14, 60% of the project managers and data analysts indicated that the frequency of disseminating information from the assessment twice in the duration of the project and the rest indicated others and they specified that they did so whenever they had information or as often as they have information. These responses went to show that the project valued the informative process and ensured information was available for most time in the duration of the project which is in line with the role of DQA as guiding the informative process.

4.4.12 Roles the Information Disseminated Plays in the Improvement of the Projects Undertaken

In this section the researcher asked the respondents to indicate the role the information disseminated plays in the improvement of the projects they undertook. In their response the project managers and data analysts indicated that from the feedback what they got determined where to focus on and know where they are going wrong, decision making, planning, evaluation and monitoring of projects, dictates the level of findings, improves on the activities of project implementation, enables them keep timelines, it helps in continuous improvements and helps in showing aspects that are missing from the project.

4.4.13 Common Ways Used to Disseminate Information

The researcher asked the respondents to indicate the most common ways used to disseminate information. In their response the project managers and data analysts indicated that they used emails to all project management staff, flyers and posters, brochures and reports at the end of the
project, data review meetings, project websites and blogs, management meetings and evaluation meetings.

4.4.14 Procedures Used to Facilitate the Process of Dissemination

In this section the researcher asked the respondents to indicate if there were any set procedures used to facilitate the process of dissemination. Their responses are shown in Table 4.15.

Table 4.11: Procedures Used to Facilitate the Process of Dissemination

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>5</td>
<td>100.0</td>
</tr>
<tr>
<td>No</td>
<td>0</td>
<td>.0</td>
</tr>
<tr>
<td>I don’t know</td>
<td>0</td>
<td>.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>73</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

From the Table 4.15, all the project managers and data analysts indicated that indeed there were procedures used to facilitate the process of dissemination.

The project managers and data analysts were also asked how these procedures enhanced and facilitated the process of dissemination. They indicated that they provide a systematic approach to the process making it more effective and efficient, regular data review writing, when a management meeting is called one is sure that all receive the reports, a follow up with emails ensures timely disseminations and they ensure that data is disseminated to the right recipients.

4.5 Standard Operating Procedures

At this point, the researcher sought to investigate how Standard Operating Procedures in DQA influence performance of HIV projects in NASCOP. The researcher asked the respondents if there were any systematic procedures that govern the process of data collection and to rate the aspects related to standard operating procedures.
4.5.1 Systematic Procedures that Govern the Process of Data Collection

In this section the researcher asked the respondents to indicate if there are any systematic procedures that govern the process of data collection. Their responses are shown in Table 4.16.

Table 4.12: Systematic Procedures that Govern the Process of Data Collection according to Data clerks, District Aids and STI Co-coordinators and Health Records Information Officers

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>70</td>
<td>95.9</td>
</tr>
<tr>
<td>No</td>
<td>1</td>
<td>1.4</td>
</tr>
<tr>
<td>I don’t know</td>
<td>2</td>
<td>2.7</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>73</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

From the Table 4.16, 95.9% of the data clerks, DASCOs and HRIOs indicated that there were systematic procedures that govern the process of data collection, 2.7% of the data clerks, DASCOs and HRIOs did not know while 1.4% of the respondents indicated that there were no systematic procedures that govern the process of data collection.

Table 4.13: Systematic Procedures that Govern the Process of Data Collection according to Project managers and Data Analysts

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>5</td>
<td>100.0</td>
</tr>
<tr>
<td>No</td>
<td>0</td>
<td>.0</td>
</tr>
<tr>
<td>I don’t know</td>
<td>0</td>
<td>.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>73</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>
From the table 4.17, all the project managers and data analysts indicated that there were set procedures designed to govern the DQA process. The project managers indicated that indeed they gave a systematic approach of the process ensuring no aspects are skipped and this leads to an improved quality data. From the above two tables, the researcher found that there were clear and entrenched operating procedures for both data collection and DQA making these process easier to conduct and efficient.

### 4.5.2 Rating of Aspects Related to Standard Operating Procedure

Here, the researcher asked the respondents to rate the following aspects related to data collection and the informative process using a likert scale of 1 to 4 with 1 being strongly agree, 2 being agree, 3 being disagree and 4 being strongly disagree. Their responses are shown in Table 4.18.

**Table 4.18: Rating of Aspects Related to Standard Operating Procedure**

<table>
<thead>
<tr>
<th>Category</th>
<th>S.A</th>
<th>A</th>
<th>D</th>
<th>S.D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>%</td>
<td>F</td>
<td>%</td>
</tr>
<tr>
<td>The procedures play a part in ensuring quality of data collected is improved and hence quality of the overall project.</td>
<td>30</td>
<td>41.1%</td>
<td>43</td>
<td>58.9%</td>
</tr>
<tr>
<td>The procedures are relevant in helping you perform your duties.</td>
<td>37</td>
<td>50.7%</td>
<td>33</td>
<td>45.2%</td>
</tr>
<tr>
<td>The project is more effective due to the procedures that are specified.</td>
<td>24</td>
<td>32.9%</td>
<td>48</td>
<td>65.8%</td>
</tr>
<tr>
<td>The project management follows up to ensure the procedures are followed as specified.</td>
<td>30</td>
<td>43.8%</td>
<td>33</td>
<td>45.2%</td>
</tr>
</tbody>
</table>

**S.A** – Strongly Agree; **A** – Agree; **D** – Disagree; **S.D** – Strongly Agree; **F** – Frequency, % - Percentage

From the Table 4.18, all the data clerks, DASCOs and HRIOs agreed with the statement that the procedures play a part in ensuring quality data collected is improved and hence quality of the overall project. 95.9% of the data clerks, DASCOs and HRIOs agreed with the statement that the
procedures are relevant in helping you perform your duties. 98.7% of the data clerks, DASCOs and HRIOs agreed with the statement that the project is more effective due to the procedures that are specified and 89.0% of the data clerks, DASCOs and HRIOs agreed with the statement that the project management follows up to ensure the procedure are followed as specified. This served to show that the project staffs were aware of the presence of the procedures and they found them pertinent in making their work easier to conduct but a number of the respondents questioned the follow-up in ensuring the procedures are followed as required.

### 4.5.3 Frequency of Performing Data Quality Assessment

The researcher asked the respondents how often they performed DQA. Their response is shown in Table 4.19

**Table 4.149: Frequency of Performing Data Quality Assessment**

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Once in the duration of the project</td>
<td>2</td>
<td>40.0</td>
</tr>
<tr>
<td>Twice in the duration of the project</td>
<td>2</td>
<td>40.0</td>
</tr>
<tr>
<td>Others</td>
<td>1</td>
<td>20.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>5</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

From the Table 4.19, 40% of the project managers and data analysts indicated that performed DQA once in the duration of the project, 40.0% of the project managers and data analysts indicated that they performed DQA twice in the duration of the project and 20.0% of the project managers and data analysts indicated that performed DQA other times specifically yearly. The findings here showed that there was a difference in the number of times projects conducted the DQA. The frequency was usually determined by the duration of the projects.
4.5.4 Skills the Staff Performing Data Quality Assessment need have to perform the Process

In this section the researcher asked the respondents to indicate how skilled the staff performing the DQA needs to be to perform this process. Their response is shown in Table 4.20.

Table 4.15: Skills the Staff Performing Data Quality Assessment need have to perform the Process

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skilled</td>
<td>5</td>
<td>100.0</td>
</tr>
<tr>
<td>Semi-skilled</td>
<td>0</td>
<td>.0</td>
</tr>
<tr>
<td>Unskilled</td>
<td>0</td>
<td>.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>5</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

From the Table 4.20, all the project managers and data analysts needed to be skilled to perform DQA. This showed that there is a high level of skill and competence in the performance of DQA and that the process required a greater amount of skills to perform.

4.5.5 Suggestions to Help in Making the Process more Effective and Efficient

In this section the researcher asked the respondents to indicate if there any suggestions they would give that would help in making the process more effective and efficient. In their response the project managers and data analysts indicated that the DQAs should be done more frequently, have constant review of indicators to make the data make sense, have the process done regularly to provide consistent data, conducting more frequent DQA for example biannually and doing it more than once to make one familiar with the process and also train more personnel.

4.5.6 If the Data Quality Assessment Process made the Impact in the Overall Performance of the Project

In this section the researcher asked the respondents to indicate if the DQA process had made an impact in the overall performance of the project. In their response the project managers and data
analysts indicated that decisions were easier to make, skilled personnel had been acquired, it had provided a targeted approach by showing what projects to dwell and focus on and also helped in the achievement of goals, it needs more support participation by all stakeholders and it had enabled them achieve most of the project objectives. The findings here went to support the fact that DQA helped in the achievement of overall objectives which in turn led to an improved performance.

4.5.7 Presence of Checklist to Facilitate the Data Quality Assessment Process

The researcher asked the respondents in conducting DQA, if the respondents made use of a checklist to facilitate the process. Their response is shown in Table 4.21

Table 4.16: Presence of Checklist to Facilitate the Data Quality Assessment Process

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>5</td>
<td>100.0</td>
</tr>
<tr>
<td>No</td>
<td>0</td>
<td>.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>5</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

From the Table 4.21, all the respondents indicated that indeed the made use of a checklist to facilitate the DQA process which was a requirement of DQA as it helped in ensuring that no aspects were missed.

The researcher asked them to further explain what was contained in this checklist and how they ensure all aspects of the checklist were met. In their response the project managers and data analysts indicated who did data abstraction, who did and how was the data analyzed and which team is in charge of validation. The checklist mainly contained the indicators of data and the quantities the data should have. It also contains what data needs to be collected, process of collection, tools for data collection and data dissemination, standardized checklist for all sites and aspects of data quality deemed important.
4.6 Data Quality Improvement

Here, the researcher sought to establish how data quality improvement in DQA influences performance of HIV projects in NASCOP. The researcher asked to indicate if the management provides any form of training to enhance your capabilities in the data collection process and to rate the aspects of improvements.

4.6.1 Training as an Improvement Mechanism

In this section the researcher asked the respondents to indicate if the management provided any form of training to enhance your capabilities in the data collection procedures. Their responses are shown in Table 4.22

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>64</td>
<td>87.7</td>
</tr>
<tr>
<td>No</td>
<td>9</td>
<td>12.3</td>
</tr>
<tr>
<td>I don’t know</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>73</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

From the Table 4.22, 87.7% of the data clerks, DASCOs and HRIOs indicated that the management provided training to enhance their capability in the data collection process while 12.3% of the data clerks, DASCOs and HRIOs indicated they did not. The researcher concluded that for the data quality to improve the staff tasked with handling it should receive some form of training to increase their capabilities to do their jobs better which in turn leads to an improved performance.
4.6.2 Rating of Aspects Related to Data Quality Improvement

The researcher asked the respondents to rate the following aspects related to data collection and the informative process using a likert scale of 1 to 4 with 1 being strongly agree, 2 being agree, 3 being disagree and 4 being strongly disagree. Their responses are shown in Table 4.23.

Table 4.18: Rating of Aspects Related to Improvement

<table>
<thead>
<tr>
<th>Category</th>
<th>S.A</th>
<th>A</th>
<th>D</th>
<th>S.D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>%</td>
<td>F</td>
<td>%</td>
</tr>
<tr>
<td>The trainings provided by the project management team are relevant to the work you do.</td>
<td>26</td>
<td>35.6%</td>
<td>43</td>
<td>58.9%</td>
</tr>
<tr>
<td>The project team is concerned in finding new ways of making the project successful.</td>
<td>18</td>
<td>24.7%</td>
<td>49</td>
<td>67.1%</td>
</tr>
<tr>
<td>Regular reviews meetings to monitor progress help in improving quality of work.</td>
<td>38</td>
<td>52.1%</td>
<td>34</td>
<td>46.6%</td>
</tr>
</tbody>
</table>

S.A – Strongly Agree; A – Agree; D – Disagree; S.D – Strongly Agree; F – Frequency, % - Percentage

From the Table 4.23, 94.5% of the data clerks, DASCOs and HRIOs agreed with the statement that the training provided by the project management team was relevant to the work they did and in turn led to an improvement of the quality of data they produced, 91.8% of the data clerks, DASCOs and HRIOs agreed with the statement that the project team is concerned in finding new ways of making the project successful which was seen as an important aspect of continuous improvement and 98.7% of the data clerks, DASCOs and HRIOs agreed with the statement that the regular reviews meetings to monitor progress helped in improving quality of work and in overall the quality of data collected and produced.
4.6.3 Way Project Management Team can Aid in Improving Process in the Project to ensure Quality Services and Achievement of Objectives

In this section the researcher asked the respondents to indicate in what other ways they thought the project management team aided in improving the process in the project to ensure quality services and the achievement of objectives. In their responses they indicated that they need to ensure that trainings are more often carried out and are job specific, increase monitoring, frequent support supervision meetings, empowering the DASCOs and HRIOs to enable data collection and site supervision, installation of electronic medical records system to ease the data processing process, provision of resources, facilitate the officers with transport when collecting the data from the respective facilities, facilitating tools and all the relevant materials in training and capacity development, through use of functional and reliable EMR across all service provision and data collection points, supervision and backstopping during the time of data collection, provision of updated tools and mentorship, reviewing the current data tools and involving the data users, modern equipment and tools to cope with the changing technology, having rotation of project staff, facilitating data collection process better, frequent meetings and seminars, capacity building in staff they work with and formulate clear communications channel to ensure correct information is shared, involving different stakeholders in development of tools, prior planning, giving feedback on the data disseminated and indicating where there is need for improvements, set out standard operating procedures and motivating staff. These responses were able to show the researcher that for there to be improvement of data quality, there needed to be continuous communication between the data clerks and project managers to find out what can be done.

4.6.4 Measures and Steps taken to Ensure that there Continuous Improvements

The researcher asked the respondents to indicate the measures and steps taken to ensure that there is continuous improvement of both the staff and process involved in DQA. In their response the project managers and data analysts indicated that the trained their staff and appraised them, removed redundant components in the process, capacity building for staff, formation of quality improvement teams, availability of data tools, regular meetings and training to ensure they are equipped with the skills.
4.6.5 Methods that Work Best in Improving the Process of Data Quality Assessment

In this section the researcher asked the respondents to indicate the methods they found to work best in terms of assisting in improving the process of DQA. In their response the project managers and data analysts indicted continuous application by the process making it easier for the staff carrying it out, functional quality improvement teams, electronic data sharing and information sharing, frequent DQAs, regular reviews of the process to add and subtract components making it more effective, utilizing the most competent staff to do DQA, collective technical reviews and qualifies data personnel.

4.7 Summary

This chapter focused on analyzing the data collected, presenting and interpreting the findings. The data was presented as per each objective and each finding was interpreted. The gender representation of the data clerks, DASCO’s and HRIO’s was equal which showed that both men and women were seen to be capable of performing the duties although the project managers and data analysts were men showing a gender disparity. The use of technology was greatly recommended so as to make the data collection and DQA more efficient. The standard operating procedures were in place but there was need to ensure they were followed to the letter. As far as the improvement was concerned, more targeted training was seen as a way of improving the quality of both the data collected and the DQA process.
CHAPTER FIVE

SUMMARY OF FINDINGS, DISCUSSION, CONCLUSIONS, RECOMMENDATIONS
AND AREAS FOR FURTHER RESEARCH

5.1 Introduction

In line with the objectives of the study, this chapter presents the summary of findings, conclusions drawn from the findings and recommendations made. The conclusions and recommendations drawn were focused on addressing the purpose of the study which was to investigate the role of data quality assessment in performance of HIV prevention projects in national AIDS and STI control programme (NASCOP) in Nyeri County.

5.2 Summary of Findings

5.2.1 Data collection and the Informative Process of Data Quality Assessment in Performance of HIV Prevention Projects

The study found that majority of the respondents agreed that data collection processes played a major role in the overall data quality hence improving performance. It found that the data collection tools were provided in adequate numbers to facilitate the process. It was noted that in as much as the respondents felt the data collection tools captured the relevant project information, there was more room for improvement. The study also found that the respondents were aware of the informative role DQA plays in project performance and the projects had several different ways of disseminating the findings. Most respondents also agreed that more still needs to be done in as far as embracing technology for the dissemination process is concerned.

5.2.2 Standard Operating Procedures of Data Quality Assessment in Performance of HIV Prevention Projects

In reference to the standard operating procedures, the study found that almost all the respondents were aware of the presence of operating procedures in performance of data collection and DQA process. It also found that the respondents found their work more efficient through the presence of these procedures. It however found that the project management was not keen in ensuring
these procedures are followed as specified. From this study it was also found that there was a checklist that was used to perform the DQA systematically.

5.3.3 Data Quality Improvement of Data Quality Assessment in Performance of HIV Prevention Projects

The study found that there were efforts of ensuring continuous improvement in the data sector to improve quality of data produced in this projects. Most respondents felt that a targeted training for the data management team would help in improving the quality of data produced.

5.3 Discussion of Findings

In a bid to answer the first research question that looked at how data collection and the informative process of DQA influence performance of HIV projects in NASCOP, it was noted that majority of the time they collected project related data monthly and other times weekly depending on the project for project undertaken for an average period of 1 years and information used to be disseminated to the stakeholders at about twice in the duration of the project. Mesfin (2006) in a study of data quality and information use in Ethiopia found that it was logical to create a sustainable framework for disseminating information even before the project has commenced. In the findings of this study we find a gap where the dissemination of information twice in the duration of a one year project as very inadequate for projects that require regular feedback and analysis as the HIV prevention projects undertaken. The projects they undertook provided them with adequate data collection tools which clearly captured the relevant information with regard to the project, and added value to the project as far as achieving project objectives concerned. This is in line with the findings of Brackstone (1999) who found that the dimensions of DQA should revolve around the ways the data is collected and the tools used to do it.

To answer the second research question that looked at how the Standard Operating Procedures on DQA influence the performance of HIV projects in NASCOP, it was noted that there were systematic procedures that govern the process of data collection process and the procedures played a part in ensuring the quality of data collected is improved, the procedures were relevant and consequently the projects were effective due to the procedures that were specified. The procedures were found to be crucial in both the data collection exercise and the DQA process.
The study found that in as much as these standards are present; they are more of imposed to the respondents. In the proceedings of the Statistics Canada Symposium (2001), it was proposed that the development of these procedures be an inclusive affair with the people carrying out the processes being involved making them more effective and efficient. On the other hand the project management supervised and made follow ups to ensure that the procedures followed were as specified. DQA processes were performed once or twice in the duration of the project to ensure that the standard operating procedures are systematic and were being followed. In line with the standard operating procedures being in place, the staffs performing DQA were required to be skilled so as to ensure the quality of data was maintained and that the projects performed as expected. Indeed there were also set procedures to govern the DQA process.

The last research questions looked at how data quality improvement in DQA help in improving performance of HIV projects in NASCOP, it was noted that management did provide training either off or on the job training to enhance the staff’s capabilities in the data collection process where it was also noted that the trainings provided by the project management was relevant. On the other hand the project teams was concerned in finding new ways of making the project success and reviewed meetings to monitor progress help in improving quality of work. Indeed it was noted there were continuous improvements of both the staff and the process involved in DQA where apart from continuous training of staff, they did appraise them in a bid to motivate them to work, removed redundant components in the process, did capacity building for staff, formed quality improvement teams, made available the data tools, had regular meetings to ensure the staff are well equipped.

5.4 Conclusions
The conclusions were drawn in line with the objectives. Based on the first objective, the study concluded that the data collected and the informative process of DQA influenced the performance of HIV projects this is so because the quality, accuracy and completeness of the data collected which was analyzed into meaningful information in deed influenced the performance of the projects because it either gave them an accurate or a tinted picture of the way things were actually on the ground.

In relation to the second objective, the study also concluded that indeed there were set standard operating procedures that governed the staff, DQA process and generally the whole process of
data collection and processing. Therefore, the standard operating procedures on DQA influenced
the performance of HIV projects in that they ensured consistency, accuracy and uniformity in
coming up with the desired results in line with project outcomes but more still needs to be done
in ensuring they are followed.

Lastly the study concluded that data quality improvement in DQA helped in improving
performance of HIV projects since information and technology is constantly evolving and new
methods of doing things are coming up every once in a while thus to reduce redundancy and
obsolescence of information that is useful to the projects. The adoption of technology should be
looked into as a tool to improve data quality.

5.5 Recommendations
In light of the above findings the researcher recommends that all the staff should be trained on
data collection and information dissemination and standard operating procedures so as to carry
out the data analysis effectively and efficiently. The staff needs to retained and appraised,
introduce capacity building for staff, formation of quality improvement teams, availability of
data tools, regular meetings and training to ensure they are equipped with the skills.

The study also recommends that there is need to install of electronic records system to ease the
data processing process, modern equipment and tools to cope with the changing technology
through the use of functional and reliable EMR across all service provision and data collection
points as this enable the staff enter data, analyze data and interpret data more effectively.

Lastly the researcher recommends that a project website should be put in place to help in
improving work quality in terms of disseminating project related information given that
nowadays internet and websites have become the commonly used channel of communication.
They also need to come up with clear communication channels to disseminate the information so
that the message is able to reach the intended audience making it useful and achieve its intended
objectives.

5.6 Areas of Further Research
Following the study, there are several areas that the researcher felt were gaps and could be
adequately pursued for further research. The Data Quality Assessment methodologies and how
they are developed and applied as well as the linkage of indicators to these methodologies in the
DQA and the effect they have on projects.

5.7 Summary
This chapter has discussed the major findings of the study based on the objectives, conclusions
drawn from the findings and recommendations to be made. Suggestions of areas for further
research are also discussed
REFERENCES


Carol S. Carson (2000), What is Data Quality, Statistics Department, International Monetary Fund.


International Monetary Fund (IMF), *Generic Data Quality Assurance Framework*, 2003


Redman, Thomas C., *Data Quality Management and Technology*, Bantam Books, New York,

Routine Data Quality Assessment Tool (RDQA), Guidelines For Implementation,


The Qualitative Report Volume 8 Number 4 December 2003 597-607
http://www.nova.edu/ssss/QR/QR8-4/golafshani.pdf


U.S. Environmental Protection Agency (EPA). 2000b. Guidance for Data Quality Assessment:


USAID, Performance Monitoring and Evaluation Tip #18 Conducting Data Quality Assessments, 2010


APPENDICES

Appendix 1: Letter of Transmittal

Valentine M. Ngumbau,

P.O. Box 110-10100,

Nyeri.

Date: .....................

To: .................................................................

Dear respondent,

RE: Letter of Transmittal of data collection

I wish to inform that I am undertaking research for my Masters of Arts degree in Project Planning and Management in the University of Nairobi. The study deals with the Role of Data Quality Assessment in Performance of HIV Prevention Projects in NASCOP, Nyeri County.

Your assistance on data collection will be appreciated as the study will benefit the entire project teams and other stakeholders. The information will be treated with confidentiality and I therefore request you to answer the questions honestly.

Attached please find the questionnaires which you are required to fill and provide information by answering the questions.

Please cooperate to enhance this process. Thank you.

Yours faithfully

Valentine M. Ngumbau

Registration No. L50/77511/2012
Appendix 2: Letter of Authority from National Commission for Science and Technology

REPUBLIC OF KENYA

NATIONAL COUNCIL FOR SCIENCE AND TECHNOLOGY

Telephone: 254-020-2213471, 2241349, 254-020-2673550
Mobile: 0713 788 787, 0735 404 245
Fax: 254-020-2213215
When replying please quote
secretary@ncst.go.ke

Our Ref: NCST/RCD/14/013/1287

Date: 16th July 2013

Valentine Mbulwa Ngumbau
University of Nairobi
P.O. Box 598
Nyeri.

RE: RESEARCH AUTHORIZATION

Following your application dated 11th July, 2013 for authority to carry out research on “Role of data quality assessment in performance of HIV prevention projects in national AIDS and STI control programme (NASCOP), Nyeri County, Kenya.” I am pleased to inform you that you have been authorized to undertake research in Othaya and Mukurweini Districts for a period ending 30th September, 2013.

You are advised to report to the District Commissioners and District Education Officers, Othaya and Mukurweini Districts before embarking on the research project.

On completion of the research, you are expected to submit two hard copies and one soft copy in pdf of the research report/thesis to our office.

DR. M. K. RUGUTI, PhD, HSC.
DEPUTY COUNCIL SECRETARY

Copy to:

The District Commissioners
The District Education Officers
Othaya District,
Mukurweini District.

“The National Council for Science and Technology is Committed to the Promotion of Science and Technology for National Development”. 

73
Appendix 3: Letter of Authority from Nyeri National AIDS and STI Control Program Office

TO WHOM IT MAY CONCERN,

RE: INTRODUCTION OF VALENTINE MBULWA NGUMBAU

This is to introduce that the above named person Valentine Mbulwa Ngumbau, ID Number 24725591, is an employee of National Aids & STI Control Program (NASCOP). She is currently undertaking a research on the “Role of Data Quality Assessment in performance of HIV Prevention Projects in NASCOP, Nyeri County” and requires your assistance with the collection of data.

Kindly accord her any assistance she may require.

Thank you.

Charles Kamau
For Provincial Aids & STI Coordinator
CENTRAL PROVINCE
Appendix 4: Letter of Introduction from University

UNIVERSITY OF NAIROBI
COLLEGE OF EDUCATION AND EXTERNAL STUDIES
SCHOOL OF CONTINUING AND DISTANCE EDUCATION
DEPARTMENT OF EXTRA MURAL STUDIES
P.O. Box 598 - NYERI Tel: 061-2030460

31 July 2013

TO WHOM IT MAY CONCERN

SUBJECT: INTRODUCTION LETTER
VALENTINE MULWA NGUMBAU - REG. NO.L50/77511/2012

This is to confirm that the above named is a bona fide student of University of Nairobi pursuing a Master of Arts Degree in Project Planning and Management - in the School of Continuing and Distance Education – Department of Extra Mural Studies.

She has completed course work and is currently writing the Research Project which is a requirement for the award of the Masters Degree.

Her topic is “Role of Data Quality Assessment in Performance of HIV Prevention Projects in NASCOP – Nyeri County- Kenya.”

Any assistance accorded to her will be highly appreciated.

UNIVERSITY OF NAIROBI
C.E.E.S. F.E.S. D.E.M.S.
NYERI & MT. KENYA AREA
P.O. BOX 535, NYERI TEL. 2966

Dr. L. Otieno - Omutoko
RESIDENT LECTURER
NYERI & MT. KENYA REGION
Appendix 5: Questionnaire for Data Clerks, DASCO’s and HRIO’s

INTRODUCTION
My name is Valentine Ngumbau and I am a student at the University of Nairobi-Nyeri Extra Mural Campus. I am undertaking a research on the Role of Data Quality Assessment in HIV prevention projects in NASCOP, Nyeri County and this questionnaire will help in data collection. Kindly answer the questions below truthfully. Please note that your response to these questions will be confidential and shall be used for the purposes of research only.

INSTRUCTIONS
Kindly tick (✓) your answer inside the box to indicate the correct answer where choices are given. Write your answer in the spaces provided where choices are not given.

Section A: Background Information

1. Please indicate your gender
   Male □ Female □

2. What is your age in years?
   18-24 □ 41-45 □
   25-30 □ 46-50 □
   30-35 □ over 50 □
   36-40 □

3. Indicate the highest educational level you have attained.
   Primary □
   Secondary □
   Certificate □
   Diploma □
   Degree and above □
4. What is your position?
   Data Clerk
   DASCO
   HRIO

Section B: Data Collection and the informative Process

5. How often do you collect project related data?
   Daily
   Weekly
   Fortnightly
   Monthly

6. Do you have a project website?
   Yes
   No
   I don’t know

Likert Scale

<table>
<thead>
<tr>
<th></th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
</table>
7. The project provides adequate data collection tools
8. The data collection tools clearly capture the relevant information with regards to the project.
9. The data collected adds value to the project as far as achieving project objectives is concerned.
10. The project management team adequately facilitates the data collection exercise?
11. The project website helps in improving work quality in terms of disseminating project related information.
12. Site discussions help improve quality of data collected and help in disseminating information.

13. Dissemination of project information is seen as a way of measuring project success.

14. What other ways are used within the project environment to ensure findings are known to all stakeholders?
   ........................................................................................................................................
   ........................................................................................................................................
   ........................................................................................................................................

15. In what ways do you think that the project management team can help in ensuring the data collection and information dissemination is more effective?
   ........................................................................................................................................
   ........................................................................................................................................
   ........................................................................................................................................

Section C: Standard Operating Procedures

16. Are there any systematic procedures that govern the process of data collection?
   
   Yes  
   No  
   I don’t know  

<table>
<thead>
<tr>
<th></th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>17.</td>
<td>The procedures play a part in ensuring quality of data collected is improved and hence quality of the overall project.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18.</td>
<td>The procedures are relevant in helping you perform your duties.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19.</td>
<td>The project is more effective due to the procedures that</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Section D: Improvement

21. Does the management provide any form of training to enhance your capabilities in the data collection process?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
<th>I don’t know</th>
</tr>
</thead>
</table>

22. The trainings provided by the project management team are relevant to the work you do.

<table>
<thead>
<tr>
<th></th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
</table>

23. The project team is concerned in finding new ways of making the project successful.

<table>
<thead>
<tr>
<th></th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
</table>

24. Regular reviews meetings to monitor progress help in improving quality of work.

<table>
<thead>
<tr>
<th></th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
</table>

25. In what other ways do you think the project management team can aid in improving process in the project to ensure quality services and achievement of objectives?

................................................................................................................................................
................................................................................................................................................
................................................................................................................................................
................................................................................................................................................
Appendix 6: Interview guide for Project Managers and Data Analysts

INTRODUCTION

My name is Valentine Ngumbau and I am a student at the University of Nairobi-Nyeri Extra Mural Campus. I am undertaking a research on the Role of Data Quality Assessment in HIV prevention projects in NASCOP, Nyeri County and this interview will help in data collection. Kindly answer the questions below truthfully. Please note that your response to these questions will be confidential and shall be used for the purposes of research only.

Section A: Background Information

1. Gender
   
   Male [ ]
   
   Female [ ]

2. What is your age in years?
   
   18 – 29 [ ]
   
   30 - 39 [ ]
   
   40 - 49 [ ]
   
   50 - 59 [ ]
   
   60 and above [ ]

3. Highest educational level attained
   
   Primary [ ]
   
   Secondary [ ]
   
   Certificate [ ]
Diploma

Degree and above

4. What is your position?

Project Manager

Data Analyst

Section B: Data Collection and Information Dissemination

5. What is the average duration of the projects you undertake?

1-4 months

4-8 months

8-12 months

Over 1 year

6. What are the most critical issues from the data you collect that contribute to data quality in these projects?

6. What are the most important components of quality data that are assessed in the projects?
8. As a project manager, how do you manage the expectations of data quality from the various data audiences?

........................................................................................................................................
........................................................................................................................................
........................................................................................................................................

9. In your opinion and from your experience, what are the most common data quality problems you have encountered?

........................................................................................................................................
........................................................................................................................................
........................................................................................................................................
........................................................................................................................................

10. How often do you disseminate information from the assessment?

   Once in the duration of the project
   Twice in the duration of the project
   Others
     If others, please specify

........................................................................................................................................

11. What role does the information disseminated play in the improvement of the projects you undertake?

........................................................................................................................................
........................................................................................................................................
........................................................................................................................................
........................................................................................................................................

12. What are the most common ways used to disseminate this information?
13. Are there any set procedures used to facilitate this process of dissemination?

   Yes   
   No    
I don’t know

   B). How do these procedures enhance and facilitate the process of dissemination?

   ………………………………………………………………… ………………………………
   …………………………………………………………………………………………………
   …………………………………………………………………………………………………

Section C: Standard Operating Procedures

14. How often do you perform DQA?

   Once in the duration of the project 
   Twice in the duration of the project 
   Others 
   If others, Please specify ………………………………………………………………… ...

15. How skilled do you and the staff performing the DQA need to be to perform this process?

   Skilled   
   Semi-skilled 
   Unskilled
16. How does this process help in ensuring the quality of data is maintained at a high and how does it help in project performance?

………………………………………………………………………………………………
………………………………………………………………………………………………
………………………………………………………………………………………………

17. Are there any suggestions you would give that would help in making the process more effective and efficient?

………………………………………………………………………………………………
………………………………………………………………………………………………
………………………………………………………………………………………………

18. In your experience, has the DQA process made an impact in the overall performance of the project?

………………………………………………………………………………………………
………………………………………………………………………………………………
………………………………………………………………………………………………

19. Are there any set procedures designed to govern the DQA process?

Yes     [ ]
No      [ ]
I don’t know     [ ]

b). In your opinion, are these procedures helpful in streamlining the process?

………………………………………………………………………………………………
………………………………………………………………………………………………
………………………………………………………………………………………………
20. In conducting DQA, do you make use of a checklist to facilitate the process?

Yes ☐

No ☐

b. Please explain what is contained in this checklist and how you ensure all aspects of the checklist are met.

………………………………………………………………………………………………
………………………………………………………………………………………………
………………………………………………………………………………………………

Section D: Improvement Process

21. What measures and steps do you take to ensure that there is continuous improvement of both the staff and process involved in DQA?

………………………………………………………………………………………………
………………………………………………………………………………………………
………………………………………………………………………………………………

22. In your experience what methods have you found to work best in terms of assisting in improving the process of DQA?

………………………………………………………………………………………………
………………………………………………………………………………………………
………………………………………………………………………………………………
……………………………………………………
# Appendix 9: Data Quality Assessment Check list

## Data Quality Assessment Checklist

<table>
<thead>
<tr>
<th>USAID Mission or Operating Unit Name:</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Title of Performance Indicator:</th>
</tr>
</thead>
</table>

* [Indicator should be copied directly from the Performance Indicator Reference Sheet]*

<table>
<thead>
<tr>
<th>Linkage to Foreign Assistance Standardized Program Structure, if applicable (i.e. Program Area, Element, etc.):</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Result This Indicator Measures [For USAID only] (i.e., Specify the Development Objective, Intermediate Result, or Project Purpose, etc.):</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Data Source(s):</th>
</tr>
</thead>
</table>

* [Information can be copied directly from the Performance Indicator Reference Sheet]*

<table>
<thead>
<tr>
<th>Partner or Contractor Who Provided the Data:</th>
</tr>
</thead>
</table>

* [It is recommended that this checklist is completed for each partner that contributes data to an indicator– it should state in the contract or grant that it is the prime’s responsibility to ensure the period for which the data are being reported]*

<table>
<thead>
<tr>
<th>Period for Which the Data Are Being Reported:</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Is This Indicator a Standard or Custom Indicator?</th>
</tr>
</thead>
<tbody>
<tr>
<td>___Standard Foreign Assistance Indicator</td>
</tr>
<tr>
<td>___Custom (created by the OU; not standard)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data Quality Assessment methodology:</th>
</tr>
</thead>
</table>

* [Describe here or attach to this checklist the methods and procedures for assessing the quality of the indicator data. E.g. Reviewing data collection procedures and documentation, interviewing those responsible for data analysis, checking a sample of the data for errors, etc.]*

<table>
<thead>
<tr>
<th>Date(s) of Assessment:</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Assessment Team Members:</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>USAID Mission/OU Verification of DQA Team Leader Officer approval</th>
</tr>
</thead>
</table>

| X |
| VALIDITY – Data should clearly and adequately represent the intended result. |
|---|---|---|
| 1 | Does the information collected measure what it is supposed to measure? (E.g. A valid measure of overall nutrition is healthy) | YES | NO | COMMENTS |
| 2 | Do results collected fall within a plausible range? | YES | NO | COMMENTS |
| 3 | Is there reasonable assurance that the data Collection methods being used do not produce systematically biased data | YES | NO | COMMENTS |
| 4 | Are sound research methods being used to collect the data? | YES | NO | COMMENTS |

| RELIABILITY – Data should reflect stable and consistent data collection processes and analysis Methods over time. |
|---|---|---|
| 1 | When the same data collection method is used To measure/observe the same thing multiple times, is the same result produced each time? | YES | NO | COMMENTS |
| 2 | Are data collection and analysis methods documented in writing and being used to ensure the same procedures are followed each | YES | NO | COMMENTS |

| TIMELINESS – Data should be available at a useful frequency, should be current, and should be timely enough to influence management decision making. |
|---|---|---|
| 1 | Are data available frequently enough to inform program management decisions? | YES | NO | COMMENTS |
| 2 | Are the data reported the most current practically available? | YES | NO | COMMENTS |
| 3 | Are the data reported as soon as possible after collection? | YES | NO | COMMENTS |

<p>| PRECISION – Data have a sufficient level of detail to permit management decision making. the the margin of error is less than the anticipated change. |</p>
<table>
<thead>
<tr>
<th></th>
<th>Question</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Is the margin of error less than the expected change being measured? (E.g. If a change of only 2% is expected and the margin of error in a survey used to collect the data is +/- 5%, then</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Has the margin of error been reported along With the data? (Only applicable to results obtained through statistical samples.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Is the data collection method/tool being used to collect the data fine-tuned or exact enough to register the expected change? (E.g. A yardstick may not be a precise enough tool to</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**INTEGRITY – Data collected should have safeguards to minimize the risk of transcription data manipulation.**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Are procedures or safeguards in place to minimize data transcription errors?</td>
</tr>
<tr>
<td>3</td>
<td>Is there independence in key data collection, management, and assessment procedures?</td>
</tr>
<tr>
<td>3</td>
<td>Are mechanisms in place to prevent Unauthorized changes to the data?</td>
</tr>
</tbody>
</table>

**SUMMARY**
Based on the assessment relative to the five standards, what is the overall conclusion regarding the Quality of the data?

Significance of limitations (if any):

Actions needed to address limitations prior to the next DQA (given level of USG control over data):

**IF NO DATA ARE AVAILABLE FOR THE INDICATOR** | **COMMENTS**
If no recent relevant data are available for this indicator, why not?

What concrete actions are now being taken to collect and report these data as soon as possible?

When will data be reported?

**Table 4.1: Data Quality Checklist: Adapted from USAID Mission: Conducting Data Quality Assessments**