THE INTERNATIONAL JOURNAL OF SCIENCE & TECHNOLEDGE

Factors Influencing Adherence to Data Protection Guidelines among Researchers at the Kenya Medical Research Institute

Kebenei Enock Kipchirchir
Student, Department of Health Management And Informatics, Kenyatta University, Kenya
Dr. George O. Otieno
Lecturer, Department of Health Management And Informatics, Kenyatta University, Kenya
Dr. Rucha Kibaara Kenneth
Lecturer, Department of Health Management And Informatics, Kenyatta University, Kenya

Abstract:
Background
A study conducted in Kenya between January and June 2014 showed that one of the important challenges facing researchers in public health data sharing is the risk to the interests of study participants. Exposure of participant's data can lead to stigmatization, loss of privacy, and unfair competition. Data loss can be through the form of intentional and unintentional 'misuse' of data. This study sought to determine the factors influencing adherence to the data protection guidelines among the researchers at KEMRI, Kenya: the national body responsible for carrying out health research in Kenya.
Objective
The study sought to determine the individual and organizational factors influencing adherence to the data protection guidelines among health researchers in KEMRI, Kenya.
Methods
This study was conducted among health researchers at the Kenya Medical Research Institute in Nairobi, Kilifi, Kisumu, and Busia Counties, Kenya. This was a quantitative cross-sectional study design involving 141 health researchers. Stratified sampling method was used to obtain the representative sample of the whole population. Questionnaires were administered to the selected KEMRI researchers. A total of nine questions extracted from the NACOSTI guidelines were asked. A respondent was considered to have adhered if he/she has agreed to all the nine questions. A p-value of <0.05 was considered statistically significant.
Results
The significant individual factors that influences adherence to the national data protection guidelines among KEMRI researchers are common forms in which data may leak to unintended persons/places (p-value of 0.04) and research stages (p-value of 0.03).
The availability of guidelines or policies on data protection within the institute is the organizational factor which highly influences adherence to data protection with a p-value of 0.01 (this shows that it is highly significant).
Institutional Ethics Review Boards (IRB) and Data Safety & Monitoring Boards (DSMBs) clearly do not play a critical role in data protection in health research with a p-value of 0.77(this shows that it is highly insignificant).
Conclusion
These results imply that both the individual and organizational factors influence adherence to the data protection guidelines among health researchers.

Keywords: *Data protection, adherence, individual factors, organizational factors, guidelines*

1. Introduction

Research data exist in different forms depending on type of research and the discipline. Data can be in the form of numbers, texts, audio, video, electronic or physical among others (Trewhella, 2014). In health, research data may be in the

form of laboratory specimens, samples, field notebooks, electronic or manual databases, clinical records, questionnaires, laboratory results, photographs, manuscripts, artefacts, and audio-visual materials among others (Trewhella, 2014).

The Article 10 of the Declaration of Helsinki states, "It is the duty of the physician in medical research to protect the life, health, privacy, and dignity of the human subject" (WMA, 2006). Article 20 states that "every precaution should be taken to respect the privacy of the subject and the confidentiality of the patient's information" (Sheikh, 2008). The "Council for International Organizations of Medical Sciences" (CIOMS) published international guidelines for medical research involving human subjects. The guidelines states that "patients have the right to expect that their physicians and other health-care professionals will hold all information about them in strict confidence and disclose it only to those who need, or have a legal right to, the information, such as other attending physicians, nurses, or other health-care workers who perform tasks related to the diagnosis and treatment of patients" (CIOMS, 2002). The "Australian Code for the Responsible Conduct of Research" states, "Researchers given access to confidential information must maintain that confidentiality." The Irish Data Protection Acts of 1988 and 2003 provide legislations that safeguards the patient information collected for medical purposes through pseudonymisation, anonymisation, protection of explicit consent, and development of Electronic Health Records (EHRs) among others (Hawkes, 2007).

The main concepts of data management/best practices that must be put into consideration in carrying out any form of research include "data ownership, gathering, storage, protection and, retention, analysis, sharing, reporting, and data destruction" (Hochstetler, 2009). Data breach refers to a situation where private and confidential data usually the ones that can identify an individual, gets to the wrong or unintended hands (Filkins and Radcliff, 2008). Data protection programs and models aim at safeguarding inadvertent data leaks.

The "Kenya Access to Information Act No. 31 of 2016" provides the principles of safeguarding personal data in different sectors. The Section 5 of the Act states, "every person has a right to privacy with respect to their personal data relating to their private and family life." The Kenya "National Commission for Science, Technology and Innovation" (NACOSTI) provides the guidelines for conducting health research in Kenya. The guideline number 15 states that the "investigator has to put in place mechanisms to protect the safety and to respect the privacy of the research subjects, and to maintain the confidentiality of the data" (NACOSTI, 2004). The KEMRI ERC has specific standards and procedures for responsible conduct of human health research in relation to data management (KEMRI, 2004).

NACOSTI is the national body with the mandate to regulate all research activities carried out in Kenya. They developed "Guidelines for Ethical Conduct of Biomedical Research Involving Human Subjects in Kenya" in the year 2004. The guidelines in the NACOSTI document borrow from the regional and international policies/guidelines for carrying out biomedical studies.

The purpose of this research was to investigate the factors influencing adherence to the data protection guidelines among the KEMRI researchers, Kenya.

2. Literature Review

2.1. Health Research and Health Research Data

According to the "Health Research Authority" (2015), the term health, clinical or medical research refers to the systematic investigation that aims at understanding the human health. The objective of such research can either be preventive or curative in nature. It is one of the significant ways of boosting people's care and treatment around the globe.

There exist various forms of health research such as clinical trials of medical devices and drugs, operational research, and qualitative studies among others. All types of research involving human participants present different level risks and benefits and risks, hence the various ethical issues that need consideration during the study approval process (Health Research Authority, 2015).

Health research can also be categorized basing on the designs such as the Randomized Controlled Trials (RCT), surveys, case control, and cohort studies among others. The research instruments used in survey studies include the questionnaires, interviews, observations, online surveys, case study, and other qualitative approaches.

The data and information from the research process are important in the improvement of human health. The overall aim of gathering and organizing data is to aid in planning, policy formulation, and decision-making in different health programmes. Ultimately, the health research data is useful in promoting global health outcomes as well as the equity.

According to the "Canadian Institutes of Health Research" abbreviated as CIHR (2016), both the researchers and the stakeholder communities acknowledge the fact that data is an essential output and input element into research as a process. Data is also used to make evidence-based decisions. There has been a global move towards the adoption of 'open' as well as the 'open' data where people around the world share data and other resources. In Canada, the CIHR is mandated to promote the effective and ethical use of health research data and information. They ensure that the available data advances knowledge, expand the available opportunities in research and improving the quality of life as well as the health products and services. CIHR has the mandate to control effective access, analysis, connection, integration, utilization, storage, dissemination, and preservation of health research data (CIHR, 2016).

2.2. National and International Guidelines for Protecting Health Research Data

According to the UNAIDS/WHO (2012) guidance point 18, "researchers have an ongoing obligation to participants to develop and implement procedures to maintain the confidentiality and security of information collected."

The Office of Research Integrity in the United States developed the "Guidelines for Responsible Data Management in Scientific Research." It is crucial for the principal investigators and the research team to understand and address data management challenges. The researcher should take into consideration the data management practices involved carrying out research. They include aspects such as data gathering, ownership, storage, safety, retention, analysis, interpretation, sharing, destruction, and data reporting among others (Coleman, and Wells, 2014).

The "National Commission for Science, Technology, and Innovation" (NACOSTI) provides the "Guidelines for Ethical Conduct of Biomedical Research Involving Human Subjects in Kenya." Notably, the fifteenth guideline states that "the relevant ethical review committee should determine whether the investigator has put in place mechanisms to protect the safety and to respect the privacy of the research subjects, and to maintain the confidentiality of the data" (NACOSTI, 2004).

2.3. Health Research Data Protection Laws and Practices

The efforts and practices aiming at protection of privacy in the health sector should be taken in a positive and collaborative manner. It should be perceived as an opportunity to advance knowledge sharing and not a hindrance or threat to research. Many rules and regulations governing the protection of health research data promote public trust, hence the willingness to continue providing personal and private information for the research purpose. The laws and policies on data protection are not hindrances to human health research, as some people contend. The rules, regulations, and guidelines provide mechanisms of information access and use for different purposes such as research (Willison, 2007).

In the year 2000, the Canadian Government established a law that governs the use of personal information called the "Protection and Electronic Documents Act" (PIPEDA). This law provided autonomy for the individual citizens to control all aspects of their personal information such as the collection, utilization, and disclosure. However, many researchers in Canada argue that this law hinders them in accessing data for health research (Tu, et al., 2004).

2.4. Factors Influencing Adherence to Data Protection Guidelines in the Health Sector

In Germany, hospitals' employees are expected to adhere to the laws and regulations on data protection in their daily work. Foth (2016) carried out a survey among hospital workers in Germany to determine the most significant factors that influence the intention of the employees to comply with the data protection guidelines as well as the variance in intention between male and female. The findings showed that the psychological factors such as employee's attitude, subjective norms and their perceived behavior control influence the adherence level significantly. It was also evident that the intention to adhere to the data protection policies is significantly different from one gender to another (Foth, 2016).

In the year 2010, a doctor from Columbia University mistakenly exposed confidential clinical information belonging to a Presbyterian patient from the New York City. The data leaked to the Internet through the shared network when the physician attempted to deactivate one of the computer servers in the facility. The "Office for Civil Rights" investigated and reported this unfortunate incident entailing a breach of patients' privacy and confidentiality. The "New York-Presbyterian Hospital" and the "Columbia University" share the data about the patients for academic research purposes (Boulton, 2014).

A qualitative study was conducted in Kenya between January and June 2014 to investigate the views of the research stakeholders on the challenges for "Public Health Research Data Sharing in Kenya." The exposure of data occurs through stigmatization, invasion of privacy, disrespecting autonomy and unfair competition. Such exposure exists through either intentional or unintentional 'misuse' of data (Jao et al., 2015).

3. Materials and Methods

3.1. Study Design

This study adopted a descriptive research design to aid in the achievement of the set objectives. The researcher chose this design due to its flexibility and expansiveness in alleviating the potential issues that may arise in the field as the questionnaires are being administered and interpreted. It uses quantitative techniques to collect, analyze and summarize data in this research (Williams, 2007).

In this study, cross-sectional study design was used and it employed mixed methods of data collection. The researcher chose cross-sectional design due to its associated low cost, minimal time, and the ability to "capture a specific point in time."

3.2. Study Location

This research was carried out at the "Kenya Medical Research Institute" (KEMRI) in Nairobi County, Kenya. Nairobi centers selected included "Eastern & Southern Africa Centre of International Parasite Control" (ESACIPAC), "Centre for Tradition Medicine and Drug Research" (CTMDR), "Centre for Biotechnology Research and Development" (CBRD), "Centre for Virus Research" (CVR), "Centre for Respiratory Disease Research" (CRDR), "Centre for Microbiology Research" (CMR), "Centre for Public Health Research" (CPHR) and "Centre for Clinical Research" (CCR).

Other centers include the "Center for Global Health Research" (CGHR) in Kisumu, "Centre of Geographical Medicine Research Coast" (CGMRC) in Kilifi, and "Centre for Infectious and Parasitic Diseases Control Research" (CIPDCR) in Busia.
The researcher chose KEMRI centers because it is one of the leading Institutions that carry out health research in Kenya. There is a need to determine the adherence to the data protection guidelines in health research. The results will help in the formulation of data management guidelines as well as the protection of rights, welfare, interests, and the privacy of human participants enrolled in research.

3.3. Study Population

The study population for this study included the scientists at the 11 research centers of the Kenya Medical Research Institute (KEMRI) located in Nairobi, Busia, Kisumu, and Kilifi Counties, Kenya.

- Inclusion criteria
- KEMRI scientists who are employed on permanent and pensionable terms.
- The study targeted the scientists who have participated in any research project involving human subjects.
- Exclusion criteria
- Researchers who seek approval from KEMRI but work outside the KEMRI centers or not affiliated to the Institute.
- Participants who did not consent to participate in this study.

3.4. Sampling Techniques for the Research Scientists

The researcher adopted the "stratified sampling method" to identify the sample. In this "probability sampling technique", the entire/whole population were the KEMRI scientists while the centers they belong form the subgroups/strata. The random selection of the final subjects was carried out in a proportionate manner from the various subgroups. The first stage was the "Probability Proportional to Size" (PPS) (Skinner, 2016).

Table 3.1. shows the distribution of the research scientists in KEMRI across the eleven centers (KEMRI Human Resource Department, 2017)

- Sample Size Determination

The Fishers' et al, 1998, formula was used to calculate the required sample size;

$$n = \frac{Z2pq}{d2}$$

Where;
n=minimum required sample size
Z=standard normal deviation at 95% CI (1.96)
p=Proportion in the target population approximated as 50% (50% was adopted because the level adherence to data protection guidelines among health researchers is unknown)
d=absolute precision, (0.05)
n=384
Since the number of KEMRI scientists is a finite population and less than 10,000, the overall sample size was determined using the "finite population correction factor."

$$NF = \frac{n}{1+(\frac{n}{N})}$$

Where n is the sample size per the fisher's et al formula above:
N is the population size, 192
Therefore;

$$NF = \frac{384}{1+(\frac{384}{192})}$$

n= 128

The figure 128 is the baseline number that the researcher intended to reach in collecting the data. However, the researcher considered an extra 10% of the sample size, which was 13 more respondents to cater for the unreturned questionnaires making a total of 141 respondents.

Cluster No.	Center	Total No of Researchers (x)	Formula (x/X)*n	Number
1.	ESACIPAC	6	(6/192)*128	4
2.	CCR	28	(28/192)*128	19
3.	CBRD	21	(21/192)*128	14
4.	CVR	27	(27/192)*128	18
5.	CTMDR	21	(21/192)*128	14
6.	CRDR	11	(11/192)*128	7
7.	CPHR	30	(30/192)*128	20
8.	CMR	21	(21/192)*128	14
9.	CGHR	17	(17/192)*128	12
10.	CIPDCR	4	(4/192)*128	3
11.	CGMR-C	6	(6/192)*128	4
Total		**192**		**128**

Table 1: Distribution of the KEMRI Research Scientists across the 11 Centers

The final participants were selected randomly from each of the 11 KEMRI centers also called strata in this study.

3.5. Data Collection Techniques

Cross-sectional data from each of the 11 KEMRI centers was collected at one specific point in time. Data was collected from a representative subset of the population of scientists at each of the KEMRI research centers. A mixture of "open-ended and closed-ended questions" were adopted in formulating questionnaires administered to the selected KEMRI scientists. The researcher visited the scientists in their respective offices within the Nairobi centers and other satellite centers. The researcher adopted both the interviewer and the self-administered questionnaires to the investigators within Nairobi Centers.

3.6. Data Management And Analysis

As the data collected was a mixture of qualitative and quantitative data, a mixed methods approach was ideal. The data analysis was done using SPSS Version 23. The research combined various descriptive statistics such as the frequency distribution, percentages, mean/averages, and standard deviation. Inferential statistics such as regression and correlation were used to determine the degree of relationship between the independent and dependent variables.

Chi- Square test of independence was used to determine if there is a significant relationship between the organizational factors, individual factors and health research data protection practices against the adherence to the national data protection guidelines in health research.

3.7. Logistical and Ethical Considerations

The authority to carry out this research was obtained from the Kenyatta University Graduate School and KU Ethics Review Committee. The proposal was then submitted to the "National Council for Science and Technology and Innovation" (NACOSTI) to get the permit to carry out the study. The researcher sought the authority to collect data from the Kenya Medial Research Institute (KEMRI) management. Lastly, the researcher sought the consent from the respondents before distributing questionnaires.

Informed consent forms in English were distributed to the potential study participants. The principles of participant's voluntariness and autonomy to make independent decision were upheld.

Participants had a right and freedom to join or withdraw from the research participation. They were not subjected to any cost for their decision to either participate or withdraw from the study.

The researcher ensured the privacy and confidentiality of the enrolled participants using the unique codes. All the information gathered in the course of this research were utilized for the intended research purposes only. The computers, laptops and other storage devices containing the confidential information were secured using strong passwords. The student kept the filled questionnaires in a safe custody until they are destroyed at end of the study.

Community considerations; the community of interest in this research includes the enrolled participants, investigators, and the affiliated organizations. The researcher involved the community in the design of this study. This research is relevant to the participants who have been enrolled in the various studies conducted at KEMRI since their personal information will be kept private and confidential. The research is also relevant to KEMRI and NACOSTI who formulate policies on health research data management. The community of selected researchers will also be consulted in the course of conducting this research. This research will also contribute to the capacity building since it will inform the training needs on health research data management within KEMRI and other research organizations. The research results will be availed to the KEMRI scientists in different forums such as publications, workshops, conferences, and seminars. The findings of this research would be presented to the Board of Examiners of Kenyatta University Department of Health Management and Informatics, Graduate School and the Post-Modern Library. These results would be disseminated to KU Ethics and Research Committee. These results would also be published for reference and presented in conferences and workshops of relevant stakeholders.

4. Results

4.1. Socio-Demographic Characteristics of Respondents

From the *Table 4.1*, the larger proportion 85(59.4%) of the respondents are males followed by 58(40.6%) females. A total of 63 (44.3%) respondents are specialized/professional graduate while a smaller proportion 3(2.1%) being the University/college diploma. Majority of the respondents 47(32.9%) lies on age group 31-40 followed closely by age group 41-50 at 42(29.4%). A larger proportion 133(93.0%) of the respondents are research officers followed by lab technologist 5(3.5%).

Characteristic	Category	Frequency (n=139)	Percentage (%)
Gender	Male	85	59.4
	Female	58	40.6
Highest Level of education	Diploma	3	2.1
	Bachelor's degree	34	23.9
	Specialized/professional graduate or post-graduate diploma	63	44.3
	PhD	42	29.6
Age group in years	< 30	29	20.3
	31 – 40	47	32.9
	41-50	42	29.4
	>50	25	17.5
Job title	Research officer	133	93
	Data manager	1	0.7
	Data analyst	1	0.7
	Student	3	2.1
	Lab technologist	5	3.5

Table 2: Socio- Demographic Characteristics of the Participants

4.2. Individual Factors Influencing Adherence to Data Protection Guidelines

The results showed that the main form in which data may be potentially leaked to the unauthorized individuals is through the use of clinical data/records followed closely with computer readable formats. Use of photographs and test responses does not in any way result to potential data leakage (*Table 4.2*)

Restricting access to the authorized persons and use of codes to conceal participant's identity are the best ways of protecting health research data. Also, data is potentially leaked to the unintended persons at the data sharing stage.

The significant individual factors that influences adherence to the national data protection guidelines among KEMRI researchers are common forms in which data may leak to unintended persons/places (p-value of 0.04) and research stages (p-value of 0.03). All the other individual factors/variables were highly insignificant.

			High Adherence Level		Low Adherence Level				
Characteristic	**Category**	**N**	**N**	**%**	**N**	**%**	**d.f.**	**X^2**	**p-value**
Research ethics & data course	Yes	98	10	10.2	88	89.8	1	1.40	0.24
	No	40	7	17.5	33	82.5			
The major threats to health research data	Theft	50	7	14	43	86	4	0.13	0.11
	Water spillage	36	0*	0*	36	100			
	Insects	15	2	13.3	13	86.7			
	Excessive heat	27	2	7.4	25	92.6			
	Unauthorized access	12	16	12.6	111	87.4			
Mechanisms for data protection	Restricting access	54	10	18.5	44	81.5	4	1.52	0.06
	Lockable cabinets	57	4	7	53	93			
	Codes	72	10	13.9	62	86.1			
	Strong passwords	67	9	13.4	58	86.6			
	Physical security	35	4	11.4	31	88.6			
The common forms in which data may leak to unintended persons/places	Field notebooks	60	9	15	51	85	7	2.87	0.04
	Filled questionnaires	76	11	14.5	65	85.5			
	Clinical data/records	60	10	16.7	50	83.3			
	Computable readable formats	69	13	18.8	56	81.2			
	Photographs	51	6	11.8	45	88.2			
	Audio visual recordings	46	7	15.2	39	84.8			
	Test responses	41	4	9.8	37	90.2			
	Slides, samples & specimens	28	4	14.3	24	85.7			
Research stages in which data may leak	Data collection	95	7	7.4	88	92.6	4	0.59	0.03
	Data analysis	41	6	14.6	35	85.4			
	Dissemination	28	4	14.3	24	85.7			
	Data sharing	83	12	14.5	71	85.5			
	Data destruction	45	3	6.7	42	93.3			

Table 3: Individual Factors Influencing Adherence to the Data Protection Guidelines
** Acknowledged As a Weakness in Chi-Square Assumptions*

4.3. Organizational Factors Influencing Adherence to the Data Protection Guidelines in Health Research

The *Table 4.3* shows that availability of guidelines or policies on data protection within the institute is the organizational factor which highly influences adherence to data protection with a p-value of 0.01 (this shows that it is highly significant).

Institutional Ethics Review Boards (IRB) and Data Safety & Monitoring Boards (DSMBs) clearly do not play a critical role in data protection in health research with a p-value of 0.77(this shows that it is highly insignificant).

			High Adherence Level		Low Adherence Level				
Characteristic	**Category**	**N**	**N**	**%**	**N**	**%**	**d.f**	**X^2**	**p-value**
Attendance of workshops on data management.	Yes	105	10	9.5	95	90.5	1	0.98	0.06
	No	37	8	21.6	29	78.4			
The Institute organized the workshop	Yes	84	9	10.7	75	89.3	1	1.23	0.27
	No	27	1	3.7	26	96.3			
Institute avails equipment to aid in data protection	Yes	101	9	8.9	92	91.1	1	1.63	0.20
	No	29	5	17.2	24	82.8			
IRBs and DSMBs play a role in data protection	Agree	73	9	12.3	64	87.7	1	1.13	0.77
	Disagree	16	9	8.3	15	91.7			
There are institutional policies for data protection	Yes	105	7	6.7	98	93.3	1	1.99	0.01
	No	23	6	26.1	17	73.9			

Table 4: Organizational Factors Influencing Adherence to the Data Protection Guidelines in Health Research

5. Discussions

5.1. Organizational Factors Influencing Adherence to Health Research Data Protection Guidelines

This study revealed that the availability of the guidelines or policies on data protection within the institute has a high influence on the adherence to data protection (p-value=0.01). It also revealed that the workshops and trainings on data management are highly significant in adherence to data protection guidelines. These findings mean that the organization play a critical role in ensuring that the employees adhere to the available guidelines for data protection. This study also revealed that Institutional Review Boards (IRB) and Data Safety & Monitoring Boards (DSMBs) are highly insignificant in terms of the roles they play in data and human participant's protection in health research (p-value=0.77).

These findings disagree with (Fabiana, 2015) who found out that all the scales of the organizational factors have a low average score. He reported that the availability of materials and equipment at the hospitals, training opportunities, and the management commitment to the safety standards would not influence the adherence to the safety procedures among the nurses.

These findings also disagree with (Neal and Sarwate, 2016) which suggest that onus is on the IRBs to safeguard data breaches. The authors went further to indicate that the researcher must take the full responsibility for the use and management of participant's data.

5.2. *Individual factors influencing adherence to health research data protection guidelines*

Majority of the respondents believe that patient's information leaks in the form of clinical data/records (p-value of 0.02) followed closely by the computer readable formats (p-value of 0.04). The findings showed that the opinions, perceptions, and attitudes of the researchers in respect to potential data leakage have an influence on the adherence to data protection guidelines. These findings were supported by Neal and Sarwate, (2016) results which revealed that 8% of data breaches in the US hospitals are as a result of the use of Electronic Health Records.

These findings also agree with the survey targeting the hospital employees in Germany which revealed that the intention to comply with data protection standards are influenced by psychological factors such as attitude, subjective norms and perceived behavior control (Foth, 2016).

Other respondents were of the opinion that restricting access to the authorized persons (p-value of 0.04) and use of codes to conceal participant's identity (p-value of 0.04) are the best ways of protecting health research data. These findings agree with (Thomson, 2011) which revealed that several workers of the Berkeley Heart Lab, California, US accessed patient's data without prior authorization and taken the critical information to the competitor in November 2011.

Majority of the respondents believed that data is potentially leaked to the unintended persons at the data sharing stage (p-value of 0.06). These findings agree with (Rathi et al., 2012) which suggests that data sharing presents ethical and social challenges in terms of the protection of the rights and dignity of participants.

5.3. Conclusion

In conclusion, there are factors influencing adherence to the data protection guidelines among health researchers in KEMRI, Kenya. Both the individual and organizational factors have an effect on the adherence to the national guidelines on data protection.

The study findings showed that the organization has a critical role to play in ensuring adherence to the data protection guidelines. The development of institutional guidelines, organization of workshops, and availing the necessary equipment will promote adherence to data protection.

The individual factors such as the views, opinions, and beliefs of the researchers influence the adherence to the data protection guidelines. These individual factors influence the practices and culture of the researchers. The information relating to the data protection guidelines that the researchers possess is attributed to the ethics courses attended and the experience in conducting research.

6. References

i. Boulton, C. (2014). Patient Data Leak Leads to Largest Health Privacy Law Settlement. The Wall Street Journal. The CIO Journal Report. Web. Accessed [10th May 2017]. http://blogs.wsj.com/cio/2014/05/09/patient-data-leak-leads-to-largest-health-privacylaw-settlement/

ii. Canadian Institutes of Health Research (CIHR). (2016). Health Research Data-Advancing Data-Intensive Research. Web. Accessed [17th May 2017]. Available at http://www.cihr-irsc.gc.ca/e/49938.html

iii. Council for International Organizations of Medical Sciences (CIOMS). (2002). International Ethical Guidelines for Biomedical Research Involving Human Subjects. Geneva, Switzerland.

iv. Coulehan, M. B., and Wells, J. F. (2014). Guidelines for Responsible Data Management in Scientific Research. The Office of Research Integrity: US Department of Health and Human Services.

v. Dembe, A. E, Partridge, J. S, and Geist, L. C. (2011). Statistical software applications used in health services research: analysis of published studies in the U.S. *BMC Health Services Research*. 2011; 11: 252.

vi. Fabiana, O. (2015). Factors influencing adherence to standard precautions among nursing Professionals in psychiatric hospitals. USP, vol.49 no.3, São Paulo, June 2015.

vii. Foth, M. (2016). Factors Influencing the Intention to Comply with Data Protection Regulations in Hospitals: Based on Gender Differences In Behaviour and Deterrence. *The European Journal of Information Systems*, Volume 25, 2016 - Issue 2.

viii. Hawkes, B. (2007). Data Protection Guidelines on research in the Health Sector. Irish Data Commisioner's Office. Health Research Authority (HRA). 2015. What is health research? London.

ix. Henley, P. et al. (2017). Development of the Good Health Research Practice course: ensuring quality across all health research in humans. *Health Research Policy and Systems*. Volume 15, Article number: 28 (2017).

x. Hochstetler, K. (2009). Research Ethics: Data Management. University of Alaska Fairbanks. Web. Accessed [9th May 2017]. Available at: http://www.uaf.edu/ori/responsible-conduct/data-management/

xi. Hochstetler, K. (2009). Research Ethics: Data Management. University of Alaska Fairbanks. Web. Accessed [12th May 2017]. Available at: http://www.uaf.edu/ori/responsible-conduct/data-management/

xii. Houser, J. (2008). *Scientific Inquiry: Precision, Reliability, and Validity: Essential Elements of Measurement in Nursing Research*. Wiley Periodicals, Inc.

xiii. Khamadi, S.I.D. (1992). Using the Library and Writing Research Proposals, Reports and Papers. Eldoret: Moi University Press.

xiv. Kenya Medical Research Institute (KEMRI). (2004). Guidelines and Standard Operating Procedures (SOPs).

xv. Kombe, F. et al. (2015). Research Stakeholders' Views on Benefits and Challenges for Public Health Research Data Sharing in Kenya: The Importance of Trust and Social Relations. PLoS One. 2015 Sep 2;10(9).

xvi. Laura, T. et al, (2014). Independent, Dependent, and Other Variables in Healthcare and Chaplaincy Research. *The Journal of Healthcare Chaplaincy*. Volume 20, Issue 4, 2014.

xvii. Laws of Kenya. (2016). Access to Information Act No. 31 of 2016. The National Council For Law Reporting.

xviii. National Commission for Science, Technology and Innovation (NACOSTI). (2004). Guidelines For Ethical Conduct of Biomedical Research, Involving Human Subjects in Kenya. Nairobi, Kenya.

xix. Neal, D. G. and Sarwate, A. D. (2016). Privacy, security, and the public health researcher in the Era of electronic health record research. *The Online Journal of Public Health Informatics*. 2016; 8(3): e207.

xx. Parker, M. and Bull, S. (2015). Sharing Public Health Research Data: Toward the Development of Ethical Data-Sharing Practice in Low- and Middle-Income Settings. *Journal of Empirical Research on Human Research Ethics*, 2015, Vol. 10(3) 217–224.

xxi. Rathi V., Dzara K., Gross C. P., Hrynaszkiewicz I., Joffe S., Krumholz H. M., Ross J. S. (2012). Sharing of clinical Trial data among trialists: A cross sectional survey. British Medical Journal, 345.

xxii. Sheikh, A. A. (2008). The Data Protection Acts 1988 and 2003: Some Implications for Public Health and Medical Research. Health Research Board (HRB). Dublin Ireland.

xxiii. Skinner, C. J. (2016). *Probability Proportional to Size (PPS) Sampling*. Wiley StatsRef: Statistics Reference Online. 1–5.

xxiv. Thomson, L. (2012). Data Breach and Encryption, Chapter 15: Health Care Data Breaches & Information Security Addressing Threats and Risks to Patient Data. American Bar Association

xxv. Tu J., Willison D., Silver F., Fang J., et al. (2004). Impracticability of Informed Consent in the Registry of the Canadian Stroke Network. New England Journal of Medicine. 2004; 350(14):1414–22.

xxvi. UNAIDS/WHO Guidance Document. (2012). Ethical Considerations in Biomedical HIV Prevention Trials.

xxvii. Williams, C. (2007). Research Methods. *The Journal of Business & Economic Research*. March 2007 Volume 5, Number 3.

xxviii. Willison, D. J. (2007). Data Protection and the Promotion of Health Research: If the Laws Are Not the Problem, Then What Is? HHealthCare Policy. 2007 Feb; 2(3): 39–43.
xxix. World Medical Association (WMA). (2006). Declaration of Helsinki.
xxx. Yamane, Taro. (1967). Statistics, an Introductory Analysis, (2nd Ed.), New York: Harper and Row.