



Sudan University of Science and Technology
College of Graduate Studies

**An Experimental Study for Correlations of Cancer
with Environmental Factors
in Sudan (1995-2013)**

**دراسة تجريبية لارتباط السرطان مع العوامل البيئية في السودان
(2013-1995)**

A thesis submitted as partial requirement for MSc in Computer of science

By:

Najat Bakhit Elbalal Ataelkarim

Supervisor:

Prof. Dr. Dieter Fritsch

October 2016

الآية

قال تعالى :

يَرْفَعِ اللَّهُ الَّذِينَ آمَنُوا مِنْكُمْ وَالَّذِينَ
أُوتُوا الْعِلْمَ دَرَجَاتٍ وَاللَّهُ بِمَا تَعْمَلُونَ
خَبِيرٌ

الآية ﴿11﴾ سورة المجادلة

صدق الله العظيم

Dedication

To my family

To my supervisor

To my teacher

To my friends

To everybody who helped me on this work

ACKNOWLEDGEMENT

I would like to express my great thanks to my god. I would also like to thank the main supervisor Prof. Dr. Dieter Fritsch for his valuable support, help and advices. The thanks are extended to all the persons for their cooperation in my research.

Abstract:

Background : Cancer is a major health issue and one of the most common diseases in the World. In Sudan, according to hospitals reports, cancer was the third leading cause of death, and accounts for 5% of all death cases in Sudan.

Objective of this study: To study the correlation between cancer incidences and environmental factors in Sudan country, from 1995-2013.

Methodology: In this experimental study design we are using statistics for bladder, brain, prostate, breast and leukemia cancer cases, analyzing data from 2010-2013. As environmental factors we hypothesize to have an impact of the agricultural density, location of communication towers, and the population census given for 1998-2013. All data is analyzed using the software packages QGIS version 2.12.3 , SPSS version 20 and R version 3.2.3. For data analysis we introduce Heatmap Analysis, Spatial Analysis Techniques, and Correlation and Regression Statistical Analysis .For the probability P we assume values of ≤ 0.05 to be statistically significant, what means the statistical level of confidence is 95%.

Results and Conclusions: The study got the following results: Dongola county in Northern state has a high Bladder cancer incidence rate (0.003444821),and Karari county in Khartoum state has high Breast (0.046873169), Leukemia (0.015023452), Brain(0.006009381),Prostate(0.015323921) cancer incidences rates. All states have a significant strong positive correlation between population and cancer incidences except Algardarif state, which has a significant weak positive correlation figure. Furthermore, we found out a significant weak (negative) association between Agriculture density and Bladder cancer ($r_s = -0.160, p - \text{value} = .003$) and Prostate cancer ($r_s = -0.190, p - \text{value} = .001$). In addition, a significant weak (positive) association between communication towers density and Breast cancer ($r_s = 0.374, p - \text{value} = .000$) could be found. Another interesting result might be a significant moderate (positive) association between communication towers and Leukemia ($r_s = 0.403, p - \text{value} = .000$) and Brain cancer ($r_s = 0.419, p - \text{value} = .000$) for all counties in Sudan .Based on the analyses we have tried to predict the numbers of cancer cases for 2014 for Blue Nile State (which is maximum 105) and Khartoum State, which is maximum 17,579.

Generally speaking, the findings from this study support the hypothesis that there is a correlation between cancer diseases and environmental factors. These findings might be useful for researchers and governmental agencies for risk assessment, regulations, and control of environmental contaminations.

المستخلص :

المقدمة : السرطان هو قضية صحية اساسية وواحد من اكثر الامراض شيوعا في العالم .في السودان وفقا لتقارير المستشفيات ,السرطان هو السبب الثالث الذي يؤدي للوفاة بمايعادل 5% من جميع الوفيات في السودان .

هدف هذه الدراسة : دراسة الارتباط بين اصابات السرطان والعوامل البيئية في دولة السودان 1995-2013.

المنهجية : في هذه الدراسة التجريبية التصميم ، استخدمنا إحصاءات عن حالات سرطان المثانة والمخ والبروستاتا والثدي و سرطان الدم من 2010-2013. كما أن العوامل البيئية التي افترضنا ان لها تأثير هي كثافة الزراعة وموقع ابراج الاتصالات ,التعداد السكاني من 1998-2013. كل البيانات تم تحليلها باستخدام حزم برامج (QGIS نسخة 2.12.3) و (SPSS نسخة 20) و (R) نسخة 3.2.3. لتحليل البيانات تم استخدام تحليل (heatmap) وتقنيات التحليل المكاني وتحليلي الارتباط والانحدار الإحصائية. واعتبرنا الاحتمالية P من $0.05 \geq$ ذات دلالة إحصائية. ما يعني أن المستوى الإحصائي للثقة هو 95%.

النتائج والخاتمة :الدراسة حصلت على النتائج التالية :محلية دنقلا في الولاية الشمالية لديها اعلى معدل اصابة بسرطان المثانة(0.003444821),محلية كرري في ولاية الخرطوم لديها اعلى معدلات اصابات بسرطان الثدي(0.046873169) و سرطان الدم(0.015023452) والمخ (0.006009381) والبروستاتا(0.015323921).كل الولايات لديها ارتباط قوي ايجابي ذو دلالة إحصائية بين السكان واصابات السرطان ما عدا ولاية القضارف لديها ارتباط ضعيف ايجابي ذو دلالة إحصائية . كما وجدنا صلة ضعيفة سلبية ذات دلالة إحصائية بين كثافة الزراعة وسرطان المثانة (0.003 , p – value = -0.160 , $r_s = -0.190$) وسرطان البروستاتا (p – value = -0.190 , $r_s = -0.160$).بالاضافة لذلك وجدت صلة ضعيفة سلبية ذات دلالة إحصائية بين كثافة ابراج الاتصالات و سرطان الثدي (p – value = 0.000 , $r_s = 0.374$).ومن النتائج المهمة وجود صلة متوسطة ايجابية ذات دلالة إحصائية بين كثافة ابراج الاتصالات وسرطان الدم (p – value = 0.403 , $r_s = 0.403$) وسرطان المخ (p – value = 0.000 , $r_s = 0.419$) لكل محليات السودان . استنادا إلى التحليل حاولنا التنبؤ باعداد حالات السرطان لعام 2014 لولاية النيل الازرق (105 وهو الحد الأقصى) وولاية الخرطوم (17,579 وهو الحد الأقصى).

بشكل عام , نتائج هذه الدراسة تدعم فرضية ان هناك ارتباط بين امراض السرطان والعوامل البيئية . قد تكون هذه النتائج مفيدة للباحثين والجهات الحكومية لتقييم المخاطر واللوائح، والحد من تلوث البيئة.

Table of Contents

Topic	Page
الإية.....	ii
Dedication.....	iii
Acknowledgement.....	iv
Abstract in English.....	v
Abstract in Arabic.....	vi
Table of Contents.....	vii
List of Tables.....	ix
List of Figures.....	x
List of Abbreviations.....	xi
Chapter one - Introduction	
1.1 Background.....	1
1.2 The Research Problem.....	2
1.3 Research Motivation.....	2
1.4 Hypothesis.....	2
1.5 Objectives	2
1.5.1 General objective.....	2
1.5.2 Specific objectives	2
1.6 Research Scope.....	3
1.7 Structure of the thesis.....	3
Chapter Two- Literature Review	
2.1 Cancer.....	4
2.1.1 Causes of Cancer.....	4
2.1.1.1 Agriculture and Cancer.....	4
2.1.1.2 Radiation and Cancer.....	5
2.1.1.2.1 Ionizing Radiation.....	5
2.1.1.2.2 Non-ionizing Radiation.....	5
2.1.1.2.3 Cell Phone Towers and Cancer.....	6
2.3 Cancer in Sudan.....	6
2.4 Cancer in World.....	7
2.5 GIS and health.....	7
2.5.1 Benefits of Using Geographic Information Systems in Health.....	7
2.5.2 Project Example of GIS and Public Health.....	8
2.6 GIS and Cancer.....	9
2.7 Previous Studies.....	9
Chapter Three-Research Methodology	
3.1 Study Idea.....	11
3.2 Study design.....	11
3.3 Data analysis.....	11
3.4 Geographical Information System(GIS).....	11
3.4.1 Components of GIS.....	12
3.4.2 GIS Sections.....	12
3.4.3 Classification of GIS Functions.....	12

Topic	Page
3.4.4 GIS Data.....	13
3.4.4.1 Vector.....	13
3.4.4.2 Raster.....	13
3.4.5 Examples of GIS software.....	13
3.5 Quantum GIS.....	13
3.5.1 Features.....	14
3.6 Heatmap analysis.....	14
3.7 Digitizing.....	15
3.7.1 Types of Digitizing in GIS.....	15
3.7.2 Georeferencing.....	15
3.8 R Software	16
3.8.1 The advantages and drawbacks of R as a GIS.....	16
3.8.2 Use R Scripts in QGIS.....	17
3.9 Statistical Package for the Social Sciences (SPSS).....	17
3.9.1 IBM SPSS Statistics Features.....	17
3.9.2 Benefits.....	18
3.10 Correlation Analysis.....	18
3.10.1 Correlation Equations.....	18
3.10.1.1 Pearson correlation coefficient.....	18
3.10.1.2 Spearman rank correlation.....	19
3.11 Regression Analysis (RA).....	19
3.11.1 Major Uses for Regression Analysis.....	19
3.11.2 Regression Analysis Methods.....	19
3.11.3 Simple Linear Regression Equation.....	19
3.12 Ethical consideration.....	20
Chapter Four - Data And Test bed	
4.1 Data Categories.....	21
4.2 Study area.....	30
Chapter Five-Results	
5.1 Results.....	26
5.2 Discussion.....	50
Chapter Six - Conclusion And Recommendation	
6.1 Conclusion.....	51
6.2 Recommendation on further research	51
References.....	52
Appendices	

List of Tables

Table No	Table Title	Page No
Table (3.1)	A summary of the relative merits of R compared with more traditional GIS software.....	16
Table (4.1)	Cancer cases of Sudan States for the Period 1995 to 2013.....	22
Table (4.2)	Population projections (in thousand) of Sudan States for the Period 1998 to 2013.....	23
Table (5.1)	cancer cases in Khartoum state at 2014.....	39
Table (5.2)	cancer cases in Blue Nile state at 2014.....	39
Table(5.3)	The correlation coefficients of cancer incidences with population by states in Sudan , 1998–2013.....	39
Table(5.4)	The correlation coefficient and P– value of prostate cancer incidence rates with agriculture density in Sudan , 2010–2013.....	40
Table(5.5)	The correlation coefficient and P– value of bladder cancer incidence rates with agriculture density in Sudan, 2010–2013.....	41
Table (5.6)	The correlation coefficient and P–value of communication towers density with leukemia cancer incidence rates in Sudan, 2010–2013...	42
Table (5.7)	The correlation coefficient and P– value of communication towers density with brain cancer incidence rates in Sudan, 2010–2013.....	43
Table (5.8)	The correlation coefficient and P–value of communication towers density with breast cancer incidence rates in Sudan, 2010–2013.....	44

List of Figures

Fig No	Figure title	Page No
Fig(4.1)	Map showing distribution of communication towers in Sudan.....	24
Fig(4.2)	Map showing heatmap analysis of agriculture in Sudan.....	24
Fig(4.3)	Map showing the counties of Sudan country.....	25
Fig(5.1)	Map showing the cancer incidence statistics in Sudan by states at period 1995–2013.....	26
Fig(5.2)	Map showing the cancer incidence statistics in Algedaref ,White Nile, Senner,Elgazira states– Sudan,at period 1995- 2013.....	27
Fig(5.3)	Map showing the brain cancer incidence rates in Sudan by county, at period 2010 –2013.....	28
Fig(5.4)	Map showing the county has very high density of brain cancer incidences in the Sudan, at period 2010–2013.....	29
Fig(5.5)	Map showing the bladder cancer incidence rates in the Sudan by county, at period 2010 –2013.....	30
Fig(5.6)	Map showing the county has very high density of bladder cancer incidences in the Sudan, at period 2010–2013.....	31
Fig(5.7)	Map showing the breast cancer incidence rates in the Sudan by county, at period 2010 –2013.....	32
Fig(5.8)	Map showing the county has very high density of breast cancer incidences in Sudan, at period 2010–2013.....	33
Fig(5.9)	Map showing the leukemia cancer incidence rates in Sudan by county, at period 2010 –2013.....	34
Fig(5.10)	Map showing the county has very high density of leukemia cancer incidences in the Sudan, at period 2010–2013.....	35
Fig(5.11)	Map showing the prostate cancer incidence rates in the Sudan by county, at period 2010 –2013.....	36
Fig(5.12)	Map showing the county has very high density of prostate cancer incidences in the Sudan, at period 2010 -2013.....	37
Fig(5.13)	Map showing a heatmap analysis of communication towers in Sudan	38
Fig(5.14)	Map showing heatmap analysis of agriculture and prostate cancer incidence rates in Sudan, 2010-2013.....	40
Fig(5.15)	Map showing heatmap analysis of agriculture and bladder cancer incidence rates in Sudan, 2010-2013.....	41
Fig(5.16)	Map showing heatmap analysis of communication towers and leukemia cancer incidence rates in Sudan,2010-2013.....	42
Fig(5.17)	Map showing heatmap analysis of communication towers and brain cancer incidence rates in Sudan, 2010-2013.....	43
Fig(5.18)	Map showing heatmap analysis of communication towers and breast cancer incidence rates in Sudan, 2010-2013.....	44

List of Abbreviations

ASCII	American Standard Code for Information Interchange
DEM	Digital Elevation Model
DNA	DeoxyriboNucleic Acid
ERDAS	Earth Resources Data Analysis System
ESRI	Environmental Systems Research Institute
GCPs	Ground Control Points
GEOTIFF	GEOgraphic Tagged Image File Format
GIS	Geographic Information System
GML	Geography Markup Language
GPS	Global Positioning System
GRASS	Geographic Resources Analysis Support System
IBM	International Business Machines
JOSM	Java Open Street Map
QGIS	Quantum Geographic information system
RA	Regression Analysis
RF	RadioFrequency
SDTS	Spatial Data Transfer Standard
SEM	Structural Equation Modeling
SPSS	Statistical Package for the Social Sciences
US	United States
USGS	United States Geological Survey
UV	UltraViolet
WHO	World Health Organization

