SciencePG
Science Publishing Group

Modelling the Sex – Specific Prevalence of Cancer Types in Mpumalanga and Eastern Cape Provincial Hospitals in South Africa

Wezile Chitha[1], John Sungwacha Nasila[1, 2, *], Zukiswa Jafta[1], Buyiswa Swartbooi[1], Siyabonga Sibulawa[1], Onke Mnyaka[1], Natasha Williams[1], Longo-Mbenza Benjamin[1, 3, 4]

[1]Wits Health Consortium, University of the Witwatersrand, Johannesburg, South Africa

[2]Department of Statistics, Walter Sisulu University, Mthatha, South Africa

[3]Department of Internal Medicine, University of Kinshasa, Kinshasa, Democratic Republic of Congo

[4]Department of Public Health, Lomo-University Research, Kinshasa, Democratic Republic of Congo

Email address:
nasila.john@gmail.com (J. S. Nasila)
*Corresponding author

Abstract: Cancer has been identified to be a major community health issue of concern to many societies. This is of particular interest when it comes to the developing South Africa. The epidemiology of cancer cases has been made known, though still under study. This research intended to understand the prevalence of different cancers and suggest preventive measures to reduce the burden of the disease and furthermore, reduce the effect of destruction to those affected in good time. The methods for data collection and overall treatment classified the study to be a cross-sectional study whose data were collected by use of a questionnaire. The questionnaire focused on variables such as counts of breast cancer, cervix cancer counts, oesophageal cancer counts and counts of other types of cancer. The analysis was analysed by use of descriptive and inferential analyses. Outcomes were well tabulated and interpreted. The results were obtained by the application of a number of methods, which were used to perform the analysis for this study. The methods were: descriptive analysis, T-test comparisons and some were complemented by error bar plots and box-plots. The following were some of the observed results for the indicated variables: Breast Cancer: Mean (201.4545), Std Dev (18.62452), 95% Ci (164.21, 238.70); Kaposi Sarcoma: Mean (29.4167), Std Dev (6.76163), 95% Ci (15.89, 42.94); Prostate Cancer: Mean (7.7500), Std Dev (.71217), 95% Ci (-1.67, 17.17); Lung Cancer: Mean (6.9167), Std Dev (.67848), 95% Ci (1.56, 12.27); Choriocarcinoma: Mean (5.3333), Std Dev (2.77434), 95% Ci (-0.22, 0.88). It is quite fitting to understand that this research as a revelation of the establishment of some very important outcomes. Of great significance, was the discovery that breast cancer among women continued to destroy the female gender in the communities where the data were collected. Results further show that cervix cancer is another cancer on the rise with a higher prevalence rate in the stated communities.

Keywords: Statistic, Women, Epidemiology, Quantitative Analysis, Prevalence

1. Background

Cancer has become one of the deadliest and silent human killers in the world. It has killed children, women of all categories, men, including cancer medical specialists and more. It has not particularly discriminated against countries, making it a serious concern by all countries. Furthermore, it has not even discriminated against gender. Both males and females have breasts [1]. This calls for the attention of medical professionals, to curb this menace. High-level credit goes to hospitals for being custodians of medical professionalism equipped with well-trained doctors and nurses whose principal objective is to provide care for the

sick and others, who may require medical and related professional engagements, including cancer [2]. Thus, medical professionals deal with both simple as well as complex medical situations resulting from cancer. Most hospitals have a hierarchical structure with well-defined roles and regulations. The distribution of power and authority depends on the placement in the hierarchy, and responsibilities are well-defined for all members [3]. This journal article intends to present a quantitative analysis and existing information on the effects caused by cancer as both a terminal disease and its prevalence in selected parts of South Africa. In addition, there will be exposure on the type and seriousness of different observed cancer types. A number of results have been observed through collected data and statistical computations. According to available data, and according to breast cancer is the most common cancer in women based on information available on the IARC and WHO database [4]. This study has also established similar information with regard to breast cancer [5].

Cancer, like any other disease, requires the collaborative approach for its management. According to all types of information management and decision-making on all levels in an organisation are interconnected [6]. Similarly, information is communicated to tactical and strategic levels if patient care needs at the operational level change, which might lead to changes in resource allocation in or between units. Within the 27 countries of the European Union (EU27), the highest female breast cancer European age-standardized mortality rates for 2008 were estimated to be in Ireland (31.1 deaths per 100 000 women) [7-13]. While the lowest was in Spain (18.4 deaths per 100 000 women), 26 Non-metastatic breast cancer is by far the most frequent cancer among women with an estimated 1.38 million new cancer cases diagnosed in 2008 (23% of all cancers). This ranks second overall (10.9% of all cancers) [14-20]. The incidence rates vary from 19.3 per 100 000 women in Eastern Africa to 89.7 per 100 000 women in Western Europe, and are high (greater than 80 per 100 000) in developed regions of the world (except Japan) and low (less than 40 per 100 000) in most of the developing regions [16, 17]. As a result, breast cancer ranks as the fifth cause of death from cancer overall (458 000 deaths), but it is still the most frequent cause of cancer death in women for both developing (269 000 deaths, 12.7% of total) and developed countries. Furthermore, in developed regions, the estimated 189 000 deaths is almost equal to the estimated number of deaths from lung cancer (188 000 deaths) [21-23]

The data for this study has been collected from South African government hospitals. Some government hospitals (such as the Nelson Mandela Academic Hospital have contributed to this cause. The data were collected on all available cancer types in the stated hospitals. Research assistants with the aid of questionnaires collected the data. The variables collected included: cancer counts of different types, counts of hormonal treatments, counts of chemotherapy treatments, the number of those who die from cancer-related causes, the number of referrals to different hospitals and others to social workers, etc. This can be observed from the accompanying analysis displayed in tables and charts.

The analysis consists of the general descriptive statistics, analysis of variance, the T-test of differences between two selected population means and correlation coefficient analysis. Charts and tables have been included for clarity. All these analyses have been supplemented with a spectrum of qualifying additional analyses, such as plots. Overall, this study intended to understand the destruction caused by cancers, the future effects of uncontrolled cancer types, and the establishment for a better solution to those affected and the determination of new approaches aimed at early detection and furthermore, to institute more preventive measures.

2. Methods

This was a cross-sectional study. A number of hospitals were involved in the collection of the data on different cancer types, making this to be an important study. The data were collected by a stratified random sampling design. The data were collected on several variables namely: cancer counts on women, cancer counts on men, use of different cancer management procedures including chemotherapy and hormonal therapy. Other variables included; counts of those who die from cancer, counts of children victims of cancer, cancer treatment by use of radiotherapy, where high-energy rays are often used to damage cancer cells and stop them from growth and multiple divisions.

The team used a questionnaire to collect the required data. The participants were drawn from patients, who were guided by either doctors or professional nurses. The researchers who designed the questionnaire were experienced doctors, professional nurses who were involved in teaching at medical schools and doctors who were participating in research work where questionnaires were used and were practicing doctors. Issues indicated in the questionnaire were well-considered and the results well-discussed. Members of the research team comprise of a multidisciplinary approach, having specialised in different fields.

The data collected were of the four scales of measurement. While some were of the nominal scale of measurement; others were ordinal; others were interval whereas others were of the ratio scale of measurement. The data were so collected that some comparisons could be performed to determine existing differences, to determine the existence of significant relationships and create other data summaries that carry sense and that could be used to make informed decisions. Plots and tables have been intentionally performed to add value to the analyses.

The participation of selected patients was an issue that had no problem both from the point of view of the study objectives and from the statistical principle of randomisation. The data were in line with the projected objectives and the randomisation idea was naturally taken care of by the random arrival of patients in a hospital. When patients leave their homes to travel to a hospital, the statistical principle of randomisation is fulfilled in the sense that no bias is involved, no body participated in asking the patient to be sick and go to hospital. It is a natural occurrence. Thus, records were completed without any influence. The research team involved

in the study meets weekly to conceptualise the research issues and discuss the data, its capturing and the analysis performed by the Statistician. The analysis is usually accompanied by comments made by the Statistician.

The calculations include simple descriptive statistics, inferential analysis and plots of relevant types to the data and to the research questions.

Comparative analyses of cancer management in the affected hospitals

Correlation analyses among pairs of selected cancer types

The following table presents the output of correlation analysis determine over the stated variables. The table contains the Pearson correlation coefficients, the p-values and the sample sizes. The computation was processed in such a way that a p-value has either one asterisk or two asterisks. The one asterisk symbolises significance at 0.05 level of significance while two asterisks show that the observed correlation is significant at 0.01 level of significance. The table holds only a few correlations for some selected pairs of variables. Significance is determined by a comparison of the observed p-value and the chosen level of significance. The rule of thumb is that if the observed p-value is smaller than the level of significance, the null hypothesis is rejected and not rejected otherwise.

Table 1. *Correlations between different cancer types in the Mpumalanga and Eastern Cape Provinces in South Africa.*

Correlations

		BREAST CANCER	KAPOSI SARCOMA	PROSTATE CANCER
BREAST CANCER	Pearson Correlation	1	.311	-.575
	Sig. (2-tailed)		.324	.050
	N	12	12	12
KAPOSI SARCOMA	Pearson Correlation	.311	1	-.327
	Sig. (2-tailed)	.324		.299
	N	12	12	12
PROSTATE CANCER	Pearson Correlation	-.575	-.327	1
	Sig. (2-tailed)	.050	.299	
	N	12	12	12
LUNG CANCER	Pearson Correlation	.383	.258	-.203
	Sig. (2-tailed)	.219	.418	.526
	N	12	12	12
CHORIOCARCINOMA	Pearson Correlation	.253	.074	-.633[*]
	Sig. (2-tailed)	.427	.818	.027
	N	12	12	12
OESOPHOGEAL CA	Pearson Correlation	.114	.324	.146
	Sig. (2-tailed)	.724	.304	.652
	N	12	12	12
LIVER CANCER	Pearson Correlation	.[a]	.[a]	.[a]
	Sig. (2-tailed)	.	.	.
	N	3	3	3
RECTAL CANCER	Pearson Correlation	-.116	-.074	.177
	Sig. (2-tailed)	.733	.828	.602
	N	11	11	11
COLON CANCER	Pearson Correlation	-.039	.009	.257
	Sig. (2-tailed)	.904	.978	.419
	N	12	12	12
CA EYE	Pearson Correlation	.655[*]	.015	-.198
	Sig. (2-tailed)	.021	.962	.538
	N	12	12	12

Correlations

		LUNG CANCER	CHORIOCARCINOMA	OESOPHOGEAL CA
BREAST CANCER	Pearson Correlation	.383	.253	.114
	Sig. (2-tailed)	.219	.427	.724
	N	12	12	12
KAPOSI SARCOMA	Pearson Correlation	.258	.074	.324
	Sig. (2-tailed)	.418	.818	.304
	N	12	12	12
PROSTATE CANCER	Pearson Correlation	-.203	-.633[*]	.146
	Sig. (2-tailed)	.526	.027	.652
	N	12	12	12
LUNG CANCER	Pearson Correlation	1	-.277	-.234
	Sig. (2-tailed)		.383	.465
	N	12	12	12
CHORIOCARCINOMA	Pearson Correlation	-.277	1	.082
	Sig. (2-tailed)	.383		.799
	N	12	12	12
OESOPHOGEAL CA	Pearson Correlation	-.234	.082	1
	Sig. (2-tailed)	.465	.799	
	N	12	12	12
LIVER CANCER	Pearson Correlation	.[a]	.[a]	.[a]
	Sig. (2-tailed)	.	.	.
	N	3	3	3

Correlations

		LUNG CANCER	CHORIOCARCINOMA	OESOPHOGEAL CA
RECTAL CANCER	Pearson Correlation	-.403	.415	.482
	Sig. (2-tailed)	.219	.204	.133
	N	11	11	11
COLON CANCER	Pearson Correlation	-.360	.168	.206
	Sig. (2-tailed)	.250	.602	.521
	N	12	12	12
CA EYE	Pearson Correlation	.135	-.075	.442
	Sig. (2-tailed)	.675	.818	.150
	N	12	12	12

Correlations

		LIVER CANCER	RECTAL CANCER	COLON CANCER
BREAST CANCER	Pearson Correlation	.ᵃ	-.116	-.039
	Sig. (2-tailed)	.	.733	.904
	N	3	11	12
KAPOSI SARCOMA	Pearson Correlation	.ᵃ	-.074	.009
	Sig. (2-tailed)	.	.828	.978
	N	3	11	12
PROSTATE CANCER	Pearson Correlation	.ᵃ	.177	.257
	Sig. (2-tailed)	.	.602	.419
	N	3	11	12
LUNG CANCER	Pearson Correlation	.ᵃ	-.403	-.360
	Sig. (2-tailed)	.	.219	.250
	N	3	11	12
CHORIOCARCINOMA	Pearson Correlation	.ᵃ	.415	.168
	Sig. (2-tailed)	.	.204	.602
	N	3	11	12
OESOPHOGEAL CA	Pearson Correlation	.ᵃ	.482	.206
	Sig. (2-tailed)	.	.133	.521
	N	3	11	12
LIVER CANCER	Pearson Correlation	.ᵃ	.ᵃ	.ᵃ
	Sig. (2-tailed)	.	.	.
	N	3	3	3
RECTAL CANCER	Pearson Correlation	.ᵃ	1	.384
	Sig. (2-tailed)	.		.243
	N	3	11	11
COLON CANCER	Pearson Correlation	.ᵃ	.384	1
	Sig. (2-tailed)	.	.243	
	N	3	11	12
CA EYE	Pearson Correlation	.ᵃ	-.027	-.032
	Sig. (2-tailed)	.	.937	.920
	N	3	11	12

Correlations

		CA EYE
BREAST CANCER	Pearson Correlation	.655*
	Sig. (2-tailed)	.021
	N	12
KAPOSI SARCOMA	Pearson Correlation	.015
	Sig. (2-tailed)	.962
	N	12
PROSTATE CANCER	Pearson Correlation	-.198
	Sig. (2-tailed)	.538
	N	12
LUNG CANCER	Pearson Correlation	.135
	Sig. (2-tailed)	.675
	N	12
CHORIOCARCINOMA	Pearson Correlation	-.075
	Sig. (2-tailed)	.818
	N	12
OESOPHOGEAL CA	Pearson Correlation	.442
	Sig. (2-tailed)	.150
	N	12
LIVER CANCER	Pearson Correlation	.ᵃ
	Sig. (2-tailed)	.
	N	3
RECTAL CANCER	Pearson Correlation	-.027
	Sig. (2-tailed)	.937
	N	11
COLON CANCER	Pearson Correlation	-.032
	Sig. (2-tailed)	.920

Correlations		CA EYE
	N	12
	Pearson Correlation	1
CA EYE	Sig. (2-tailed)	
	N	12

Oncology comparative Data Analyses

Table 2. *Descriptive data analysis.*

Descriptive Statistics	N	Mean	Std. Deviation	95% confidence interval	
	Statistic	Statistic	Statistic	lower limit	Upper limit
BREAST_CANCER	11	201.4545	18.62452	164.21	238.70
KAPOSI_SARCOMA	12	29.4167	6.76163	15.89	42.94
PROSTATE_CANCER	12	7.7500	4.71217	-1.67	17.17
LUNG_CANCER	12	6.9167	2.67848	1.56	12.27
CHORIOCARCINOMA	12	5.3333	2.77434	-0.22	10.88
OESOPHOGEAL_CA	12	9.3333	5.03322	-0.73	19.40
LIVER_CANCER	3	1.0000	.00000	1.00	1.00
RECTAL_CANCER	11	3.8182	1.99089	-0.16	7.80
LARYNX_CANCER	7	1.1429	.37796	0.39	1.90
COLON_CANCER	12	9.1667	3.18614	2.79	15.54
CA_EYE	12	3.0000	1.75810	-0.52	6.52
PELVIS	2	1.0000	.00000	1.00	1.00
SCC_CANCER	9	2.1111	1.36423	-0.62	4.84
STOMACH_CANCER	10	1.9000	.87560	0.15	3.65
CERVIX_CANCER	12	101.5833	22.07614	57.43	145.74
CA_OVARY	12	5.3333	3.05505	-0.78	11.44
MULTIPE_MYELOM_CA	12	4.3333	1.66969	0.99	7.67
MELANOMA_CA	7	2.4286	1.81265	-1.20	6.05
ADENOCARCINOMA	10	2.2000	1.22927	-0.26	4.66
ENDOMETRIAL_CA	12	3.4167	2.53909	-1.66	8.49
LYMPHOMA_CANCER	12	11.9167	5.21833	1.49	22.34
GASTRIC	11	2.5455	1.63485	-0.72	5.82
MERKELL_CARCINOMA	6	3.8333	2.48328	-1.13	8.80
CA_ANUS	7	1.5714	.78680	0.00	3.15
CASTLEMENS_DISEASE	6	1.0000	.00000	1.00	1.00
CA_ABDOMIN	2	2.0000	1.41421	-0.83	4.83
GESTATIONAL	4	1.2500	.50000	0.25	2.25
PAREAGANGALLOMA	3	2.3333	2.30940	-2.29	6.95
BLADDER	11	3.9091	2.50817	-1.11	8.93
SINONASAL	2	3.0000	.00000	3.00	3.00
BUCCAL_TUMAR	3	1.3333	.57735	0.18	2.49
VULVA	12	3.6667	1.66969	0.33	7.01
COLORECTAL	5	1.6000	.54772	0.50	2.70
CONJUCTIVIA	7	2.0000	1.00000	0.00	4.00
CA_UTERUS	5	1.6000	.89443	-0.19	3.39
CA_THYROID	8	2.6250	1.18773	0.25	5.00
MYCOSIS_FUNGOIDES	7	2.1429	1.34519	-0.55	4.83
CHONDROBLASTOSARCOMA	1	1.0000	.	xxxx	xxxx
RHABDOMYOSARCOMA	6	1.6667	1.03280	-0.40	3.73
CA_AXILLAR	2	1.0000	.00000	1.00	1.00
HCC	9	1.8889	1.61589	-1.34	5.12
FLOOR_OF_MOUTH	6	1.5000	.83666	-0.17	3.17
CA_PANCREAS	5	1.2000	.44721	0.31	2.09
CA_BASAL	2	1.0000	.00000	1.00	1.00
OSTEOSARCOMA	7	1.2857	.48795	0.31	2.26
MALIGNANT_NEOPLASM	3	2.3333	.57735	1.18	3.49
SERTOLI_LEYDIG_TUMOR_CELL	2	1.0000	.00000	1.00	1.00
CA_ELLEOCACCAEL	3	1.3333	.57735	0.18	2.49
CA_GUM	1	1.0000	.	xxxx	xxxx
SARCOMA	6	1.8333	.98319	-0.13	3.80
OROPHARYNX	3	1.0000	.00000	1.00	1.00
BRAIN_TUMOR	1	1.0000	.	xxxx	xxxx

Two-variable periodical comparisons of cancer prevalence

This study chose to make two-variable comparisons to understand the existence of any difference between the selected variables with regard to cancer. The following comparisons were made between two adjacent months and further, between the two-recorded genders. The comparisons were made by use of the independent T-test statistic, which compares *means* of the two selected populations. The independent T-test uses the logic that the comparison is focused on different gender means within the same period. The time in space comparison is to understand the influence of time specific and gender difference on the prevalence of

cancer. The T-test performs this test by use of the following formula by which a t-test statistic is determined.

$$ t = \frac{\bar{X}_1 + \bar{X}_2}{\sqrt{S^2(\frac{1}{n_1} + n_2)}} $$

Where; t is the test statistic;

S^2 is the pooled observed sample variance;

$\bar{X}_1$ and $\bar{X}_2$ are the means of samples drawn from the two populations for comparison;

n_1 and n_2 are two sample sizes for data drawn from the two populations.

Here t follows the T-distribution with $n_1 + n_2 - 2$ degrees of freedom.

The practical comparison is based on the use of either the t test-statistic or the observed p-value. These are compared to the tabulated t-value or to the suggested level of significance, depending on the choice of the test.

As for the present comparisons, the researchers used the p-value to compare with the level of significance (0.05). Thus, an observed mean difference will be significant if the calculated p-value is smaller than the level of significant, leading to the rejection of the null hypothesis. However, if the Table 3.

Table 3. The oobserved p-value and t-test statistics according to pairs of months.

T-Test for comparison based on gender Group Statistics	Months for comparison	N	Mean	Std. Error Mean
April 2019 to May 2o19counts	Males	11	121.64	45.634
	Females	11	115.45	42.418
June 2019 to July 2019 counts	Males	11	113.73	42.499
	Females	11	140.82	50.358
August 2019 to Sept 2019 counts	Males	11	115.00	44.350
	Females	11	113.18	44.237
Oct 2019 to Nov 2019counts	Males	11	132.09	49.883
	Females	11	130.18	46.695
Dec 2019 to Jan 2020counts	Males	11	88.36	38.004
	Females	11	112.91	40.787
Feb 2020 counts	Males	11	113.18	46.057
	Females	11	128.27	47.653

		t-test for Equality of Means	
		Sig.	t
April 2019 to May 2o19counts	Equal variances assumed	.749	.099
	Equal variances not assumed		.099
June 2019 to July 2019 counts	Equal variances assumed	.530	-.411
	Equal variances not assumed		-.411
August 2019 to Sept 2019 counts	Equal variances assumed	.998	.029
	Equal variances not assumed		.029
Oct 2019 to Nov 2019counts	Equal variances assumed	.750	.028
	Equal variances not assumed		.028
Dec 2019 to Jan 2020counts	Equal variances assumed	.778	-.440
	Equal variances not assumed		-.440
Feb 2020 counts	Equal variances assumed	.837	-.228
	Equal variances not assumed		-.228

Independent Samples Test

		t-test for Equality of Means		
		df	Sig. (2-tailed)	Mean Difference
April 2019 to May 2o19counts	Equal variances assumed	20	.922	6.182
	Equal variances not assumed	19.894	.922	6.182
June 2019 to July 2019 counts	Equal variances assumed	20	.685	-27.091
	Equal variances not assumed	19.451	.685	-27.091
August 2019 to Sept 2019 counts	Equal variances assumed	20	.977	1.818
	Equal variances not assumed	20.000	.977	1.818
Oct 2019 to Nov 2019counts	Equal variances assumed	20	.978	1.909
	Equal variances not assumed	19.913	.978	1.909
Dec 2019 to Jan 2020counts	Equal variances assumed	20	.664	-24.545
	Equal variances not assumed	19.901	.664	-24.545
Feb 2020 counts	Equal variances assumed	20	.822	-15.091
	Equal variances not assumed	19.977	.822	-15.091

Independent Samples Test

		t-test for Equality of Means	
		Std. Error Difference	95% Confidence Interval of the Difference
			Lower
April 2019 to May 2o19counts	Equal variances assumed	62.304	-123.782
	Equal variances not assumed	62.304	-123.826
June 2019 to July 2019 counts	Equal variances assumed	65.895	-164.546
	Equal variances not assumed	65.895	-164.795
August 2019 to Sept 2019 counts	Equal variances assumed	62.640	-128.847
	Equal variances not assumed	62.640	-128.847
Oct 2019 to Nov 2019counts	Equal variances assumed	68.328	-140.621
	Equal variances not assumed	68.328	-140.660
Dec 2019 to Jan 2020counts	Equal variances assumed	55.749	-140.836
	Equal variances not assumed	55.749	-140.873
Feb 2020 counts	Equal variances assumed	66.272	-153.332
	Equal variances not assumed	66.272	-153.343

Independent Samples Test

		t-test for Equality of Means
		95% Confidence Interval of the Difference
		Upper
April 2019 to May 2o19 counts	Equal variances assumed	136.145
	Equal variances not assumed	136.190
June 2019 to July 2019 counts	Equal variances assumed	110.364
	Equal variances not assumed	110.613
August 2019 to Sept 2019 counts	Equal variances assumed	132.483
	Equal variances not assumed	132.483
Oct 2019 to Nov 2019 counts	Equal variances assumed	144.439
	Equal variances not assumed	144.479
Dec 2019 to Jan 2020 counts	Equal variances assumed	91.745
	Equal variances not assumed	91.782
Feb 2020 counts	Equal variances assumed	123.151
	Equal variances not assumed	123.161

3. Results

3.1. Charts Used to Compare Population Means Using Error Bars

The chart below compares cancer prevalence in the months of April 2019 and May 2019. It is observed from the error bars that the April mean of 122 is higher than that of May with a mean of 115. The difference is 7. A t-test will determine whether a difference of 7 is significant or not.
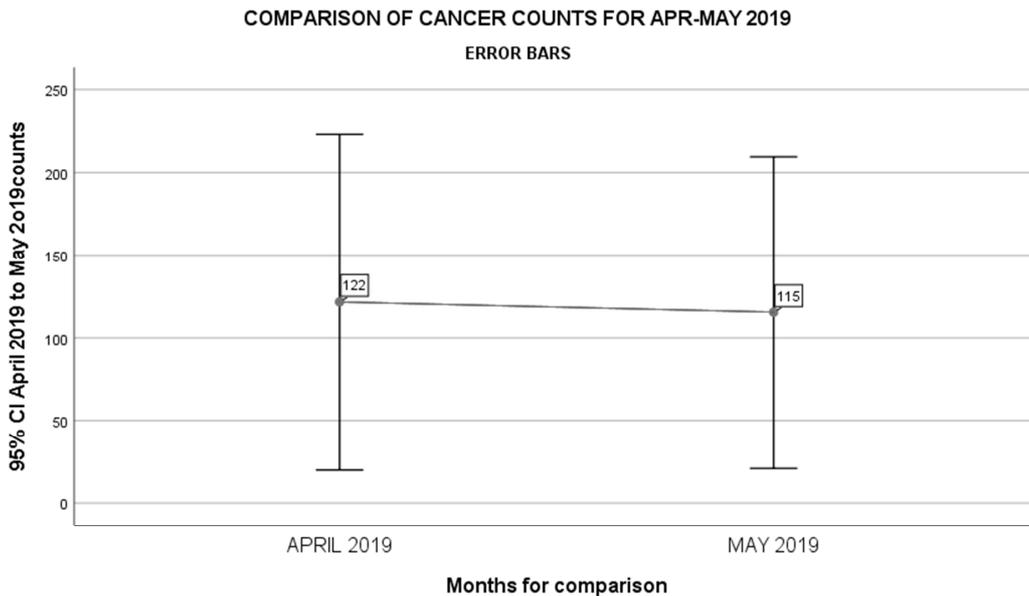

Figure 1. *April 2019 to May 2o19 counts compared using Box-plots.*

The following figure presents a chart, which compares cancer prevalence in the months of April and May 2019. In the chart, there are box-plots, which show the median values for the two months. A direct observation notices that the May 2019 median is higher than the April 2019 median.
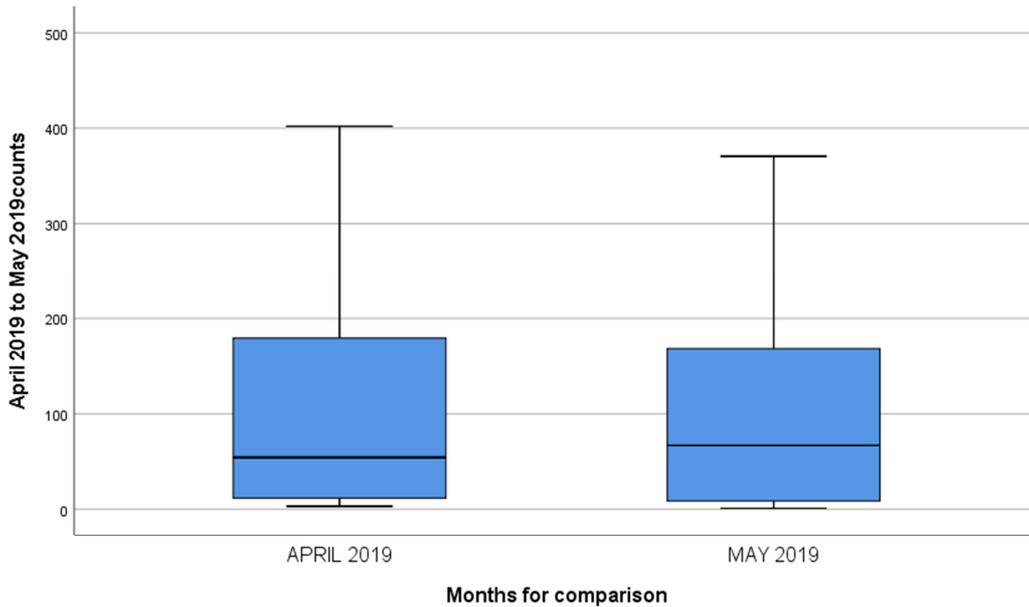

Figure 2. Chart used to compare population means using error bars for June and July 2019.

The chart below compares cancer prevalence in the months of June 2019 and July 2019. It is observed from the error bars that July with a mean of 141 is higher than that of June with a mean of 114. The difference is 27. It remains to use the T-test to understand the significance of the difference of 27 under the prevailing conditions.
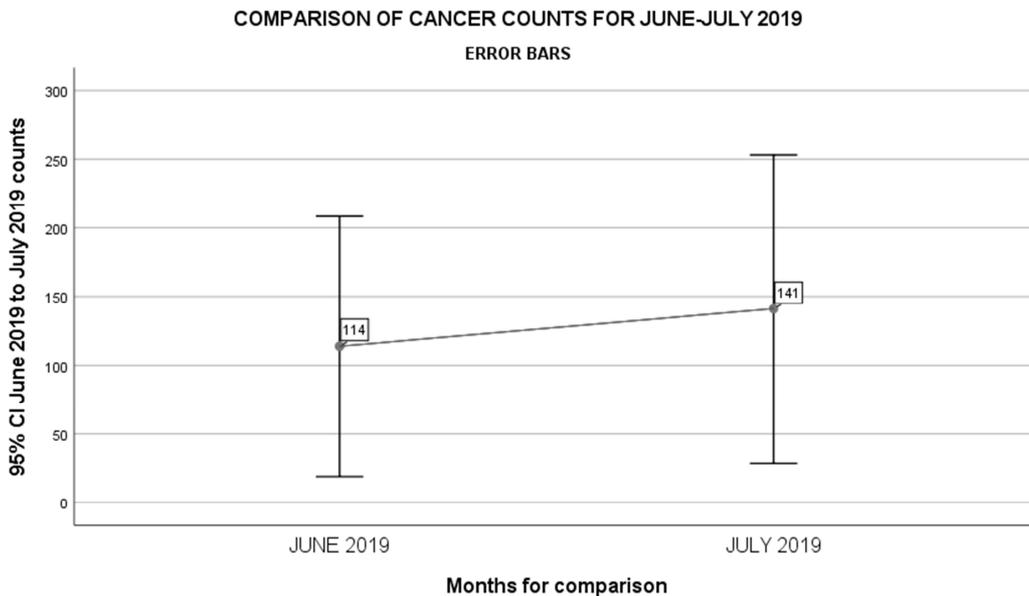
Figure 3. Figure showing the June 2019 and July 2019 counts.

The following figure presents a chart, which compares cancer prevalence in the months of June and July 2019. In the chart, there are box-plots, which show the median values of the two months. A direct observation notices that the July 2019 median is higher than the June 2019 median. These box-boxes tell us that the July 2019 data were more dispersed than the June data. The whisker distances from the median are compared to arrive at this conclusion.
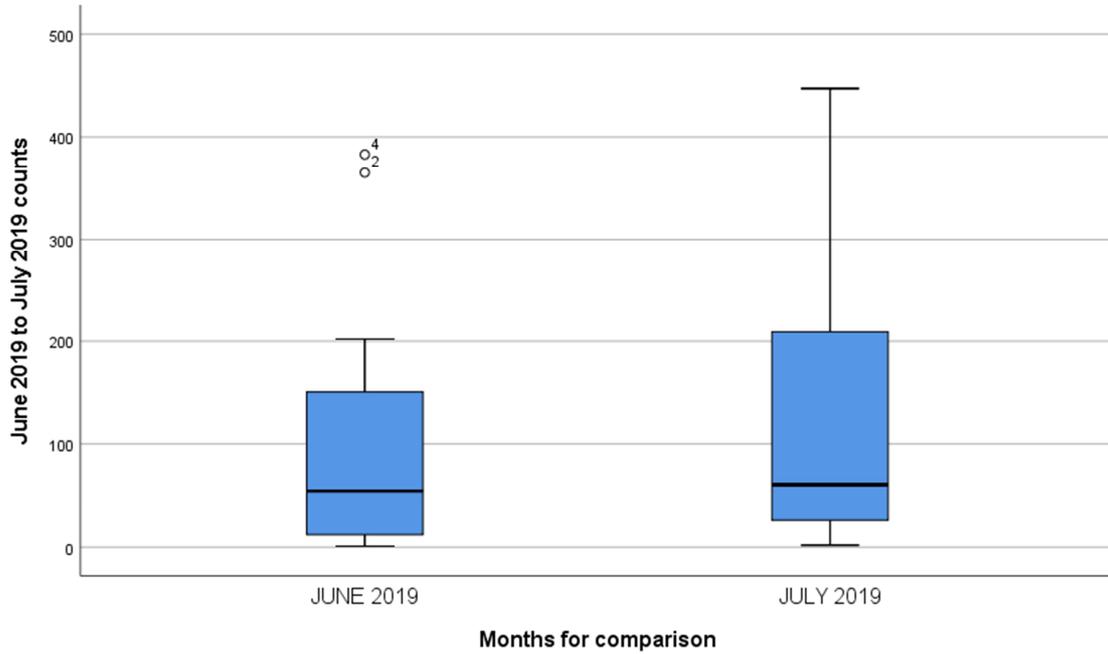
***Figure 4.** Box plot showing a comparison between counts of cancer between June and July 2019.*

The chart below compares cancer prevalence in the months of August 2019 and September 2019. It is observed from the error bars that the August mean of 115 is higher than that of September with a mean of 113. The difference is 2. It can be shown that a mean difference of 2 under the conditions of this test cannot be significant.
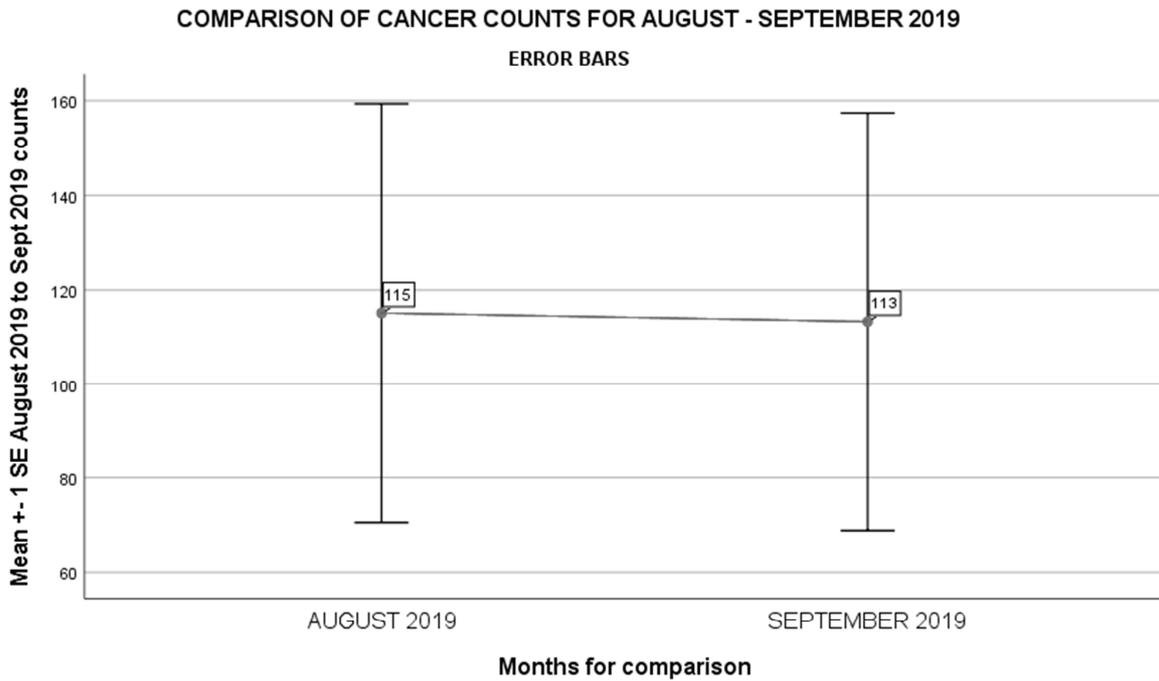

***Figure 5.** Error bar plots used to compare cancer counts for October and November 2019.*

The following error bar plot presents a comparison between the cancer counts for October and November 2019. The average count for October is observed to be 132, while that for December is 130. The difference between the two months is just two. There is a high possibility that the difference between the two months is not significant.

COMPARISON OF CANCER COUNTS FOR OCTOBER - NOVEMBER 2019

ERROR BARS
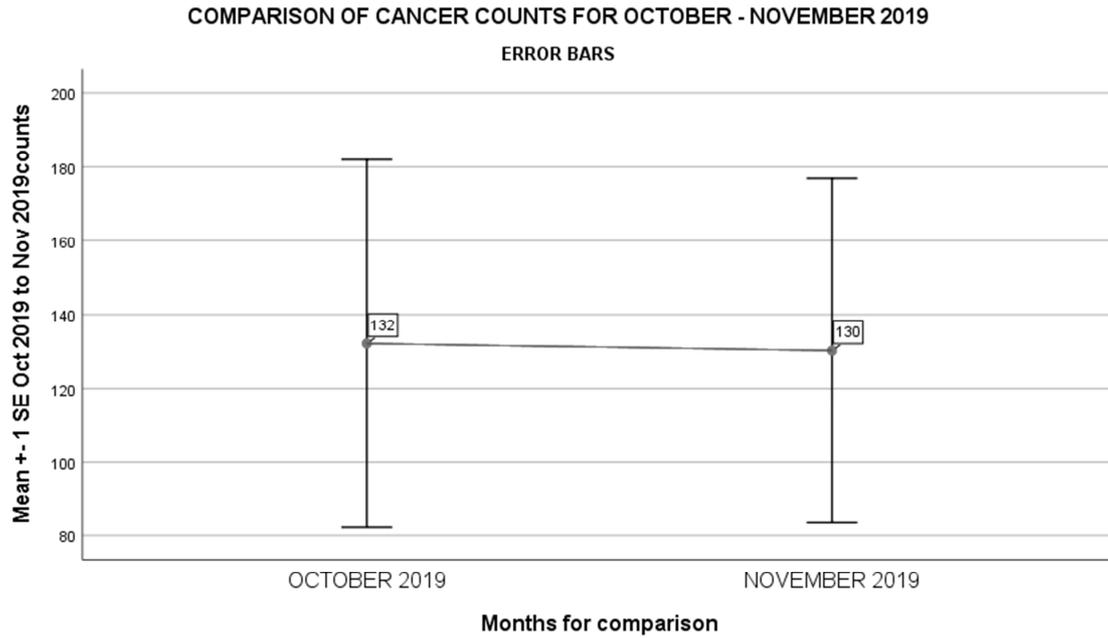
***Figure 6.** Comparison between October and December 2019 counts.*

The following box-plot compares quantile position averages over the two months of October and December. Both the box-plots show right-skewedness for each of the two months. The difference is that though they are both right-skewed, the spread for October data was more than that of November data.
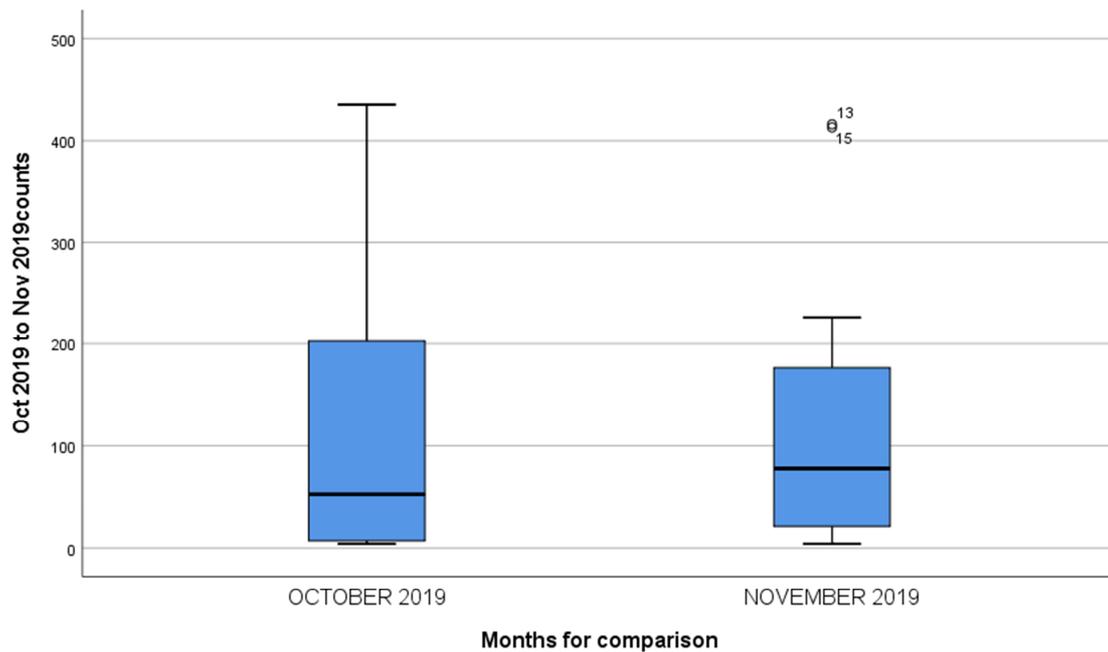

***Figure 7.** Comparison between December 2019 and January 2020 counts using box-plots.*

The following box-plot compares quantile position averages over the two months of December 2019 and January 2020. Both the box-plots demonstrate right-skewedness for each of the two months. The difference is that though they are both right-skewed, the spread for January data was far more than that of the December data. It can be observed from the plot that the median for December was 15 while that for Januarys was 60 counts. It is understood that individual quartile variations depend on the month of data count. While January 2020 showed a higher degree of variability over-all, the December data analysis shows a less significant variability. However, both of the two months have an advantage over the other separately. January has a high maximum observation whereas December has a lower variance.
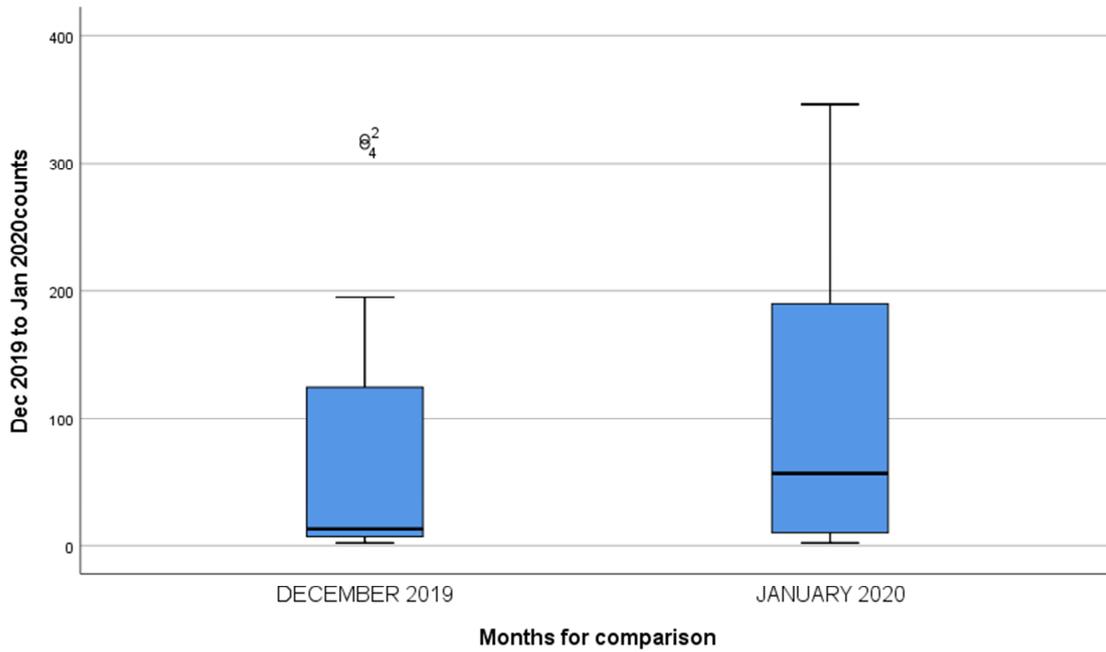
Figure 8. *Comparison between February and March 2020 counts.*

The following box-plot compares the quantiles of for the months of February and March. The figure has a lot in common with the above figure. This means that the interpretation will be of the same form and approach. Thus, it can be observed from the plot that the median for February 2020 is 30 while that for March 2020 is 60 counts. It is an understanding that individual quartile variations depend on the month of data count. While March 2020 shows a higher degree of variability over-all, the February data analysis shows a less significant variability. However, both of the two months have an advantage over the other separately. March has a higher maximum observation whereas February has a lower variance.
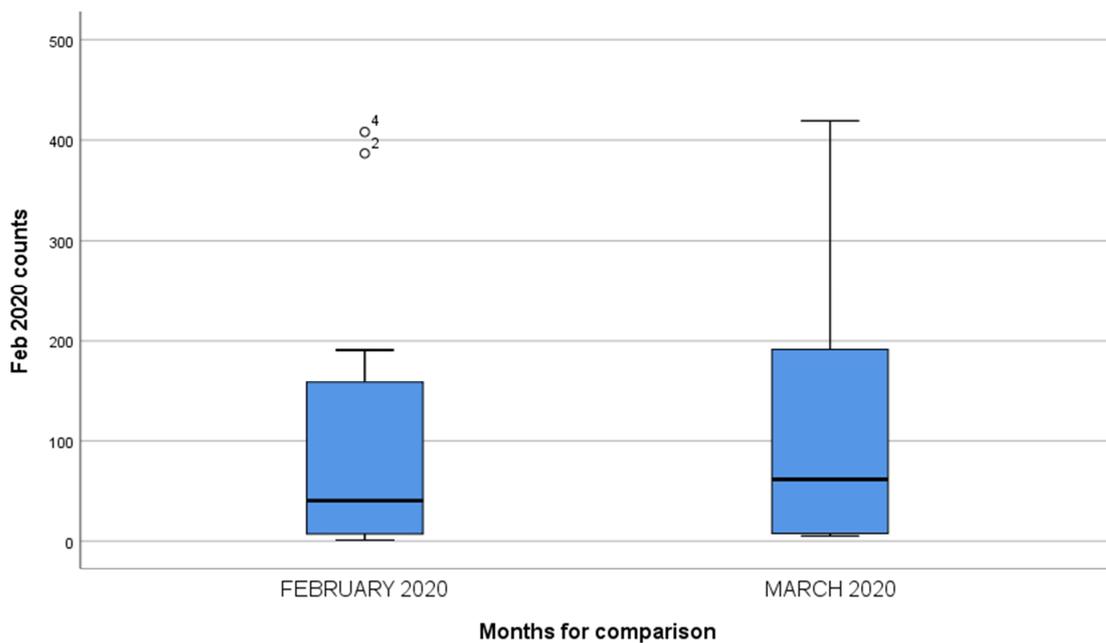

Figure 9. *Comparison between February and March 2020 counts.*

4. Discussion

This study was described to be exploratory. One of the main objectives of this study was stated earlier to be a comparison of the prevalence of the different cancer types. This knowledge would help in the promotion of a deeper understanding of the individual cancer types. The prioritisation of a more advanced understanding of the destruction caused by different cancer types will be decided by a direct comparison of the observed statistics from the analysis. In the table below, breast cancer is seen to be the most prevalent with an observed mean of 201.46, a standard deviation of 18.63 and an estimated 95% confidence interval of (164.21, 238.70). The second most important cancer is observed to be Cervix cancer, which

averaged 101.58 with a standard deviation of 22.08. This type of cancer had a 95% confidence interval estimate of (57.43, 145.74). The third cancer type in the order of decreasing average was Kaposi Sarcoma, which had a mean statistic of 29.42, a standard deviation of 6.76 and 95% confidence interval estimate of (15.89, 42.94). Other statistics can easily be read from the table below. These findings have been supported by who claims that breast cancer and cervix cancer have the highest contributions to cancer among women. It is documented, however, that oesophageal cancer is least prevalent [5]. This claim has further been proved by this research. The high level of prevalence of breast cancer and followed by cervix cancer is strongly supported by, who claimed that breast cancer is the most common cancer, which mostly affected women [1].

The following box-plot compares quantile position averages over the two months of October and December. Both the box-plots show right-skewedness for each of the two months. The difference is that though they are both right-skewed, the spread for October data was more than that of November data. It can be observed from the plot that the median for October is 50 while that for November is 70 counts. The individual quartile variations depend on the month of data count. While November shows a higher degree of consistency, the October data analysis shows a more significant variability. Both the months have an advantage over the other.

5. Conclusion

This research has established some very important outcomes. Of great significance, was the discovery that breast cancer to women continued to be destructive to women in the community where the data were collected. Another established cancer type is cervix cancer, which was ranked second to breast cancer. Breast cancer has affected men as well, though the data collected did not provide the statistical opportunity to establish good comparative results. Different treatments have been compared. Most inferential analysis using the T-test over gender and period have shown no significance at the 0.05 level of significance. The included error plots and box-plots have further confirmed this. It has been noticed that due to emerging questions, the questionnaire has been reconstructed to include other variables of importance.

Author's Contributions

WC, JSN and ZJ designed and analyzed the statistical data for the study. BS, SS, OM, NW and LMB supervised the study. All authors have read and approved the final and revised version of the manuscript.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

References

[1] Balekouzou, A., Ramatika, C., Bishwajit, G., Nambei, S., & et al. (2016). The Epidemiology of breast cancer: A retrospective study in the Central African Republic. *BMC Public Health*, 1. doi: 10.1186/s 12889-016-3863-6.

[2] Bhutan InfoComm and Media Authority, Royal Government of Bhutan. Rules on Content. Thimphu, Bhutan InfoComm and Media Authority 2010.

[3] Curado, M. Breast cancer in the world: Incidence and mortality. *Salud Publica De Mexico, 2014; 53* (5), 372-384.

[4] Lundgrén-Laine, H., & Salanterä, S. Information systems in hospitals: A review article from a nursing management perspective. *International Journal of Networking and Virtual Organisations 2013*. doi: 10.1504/IJNVO.2013.058441.

[5] Mintzberg, H. (2002). Managing care and cure-up and down, in and out. *Health Services Management Research, 15* (3), 193 - 2006.

[6] Sithole, N., & Somdyala, N. Hospital-Based Cancer Registry Frere Hospital, East London cancer Incidence 1991-209, Technical Report 2017; 10.

[7] Virtanen, J., & Kovalainen. Doctor's and nurse's organizations from the point of view of practical leadership (in Finnish: Lääkärien ja hoitajien linjaorganisaatiot käytännön johtamisen näkökulmasta). *Finnish Medical Journal, 2006; 61* (33).

[8] Ferlay J, Colombet M, Soerjomataram I, Dyba T, Randi G, Bettio M et al. Cancer incidence and mortality patterns in Europe: Estimates for 40 countries and 25 major cancers in 2018. European Journal of Cancer, 2013; 49 (6): 1374-1403.

[9] International Agency for Research on Cancer. Latest global cancer data: Cancer burden rises to 19.3 million new cases and 10.0 million cancer deaths in 2020. Press release 292. 15 December 2020.

[10] Ferlay J, Ervik M, Lam F, Colombet M, Mery L, Piñeros M, Znaor A, Soerjomataram I, Bray F (2020). Global Cancer Observatory: Cancer Today. Lyon, France: International Agency for Research on Cancer. Available from: https://gco.iarc.fr/today, accessed [28 January 2021]. *Statistics for Europe are based on WHO Europe region of 53 countries. Data for Andorra, Monaco and San Marino are not included*.

[11] European Commission. Breast cancer burden in EU-27. European Cancer Information System. https://ecis.jrc.ec.europa.eu/pdf/Breast_cancer_factsheet-Dec_2020.pdf. Accessed 28 January 2021.

[12] European Commission. European Cancer Information System. https://ecis.jrc.ec.europa.eu. Accessed 28 January 2021.

[13] Wild CP, Weiderpass E, Stewart BW (eds.). World Cancer Report: Cancer Research for Cancer Prevention. World Health Organization, Geneva 2020.

[14] Bray F, Jemal A, Grey N, Ferlay J, Forman D. Global cancer transitions according to the human development index (2008-2030): a population-based study. Lancet Oncol. 2012; 13: 790–801. doi: 10.1016/S1470-2045(12)70211-5.

[15] Fitzmaurice C, Allen C, Barber RM, Barregard L, Bhutta ZA, Brenner H, et al. Global, regional, and National Cancer Incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 Cancer groups, 1990 to 2015. JAMA Oncol. 2017; 3: 524. doi: 10.1001/jamaoncol.2017.1747.

[16] Althuis MD, Dozier JM, Anderson WF, Devesa SS. Brinton L a. Global trends in breast cancer incidence and mortality 1973-1997. Int J Epidemiol. 2005; 34: 405–412. doi: 10.1093/ije/dyh414.

[17]   Galceran J, Ameijide A, Carulla M, Mateos A, Quirós J, Rojas M, et al. Cancer incidence in Spain, 2015. Clin Transl Oncol. 2017; 19: 799–825. doi: 10.1007/s12094-016-1607-9.

[18]   Héry C, Ferlay J, Boniol M, Autier P. Quantification of changes in breast cancer incidence and mortality since 1990 in 35 countries with Caucasian-majority populations. Ann Oncol. 2008; 19: 1187–1194. doi: 10.1093/annonc/mdn025.

[19]   Hankinson SE, Colditz GA, Willett WC. Towards an integrated model for breast cancer etiology: the lifelong interplay of genes, lifestyle, and hormones. Breast Cancer Res. 2004; 6: 213–8.

[20]   Lacey J, Devesa S, Brinton L. Recent trends in breast cancer incidence and mortality. Environ Mol Mutagen 2002; 88: 82–8.

[21]   Colonna M, Delafosse P, Uhry Z, Poncet F, Arveux P, Molinie F, et al. Is breast cancer incidence increasing among young women? An analysis of the trend in France for the period 1983-2002. Breast. 2008; 17: 289–292. doi: 10.1016/j.breast.2007.10.017.

[22]   Leclère B, Molinie F, Tretarre B, Stracci F, Daubisse-Marliac L, Colonna M, et al. Trends in incidence of breast cancer among women under 40 in seven European countries: a GRELL cooperative study. Cancer Epidemiol. 2013; 37: 544–549. doi: 10.1016/j.canep.2013.05.001.

[23]   Pollán M. Epidemiology of breast cancer in young women. Breast Cancer Res Treat. 2010; 123: 3–6. doi: 10.1007/s10549-010-1098-2.