SPATIAL STATISTICAL ANALYSIS OF FEMALE BREAST CANCER INCIDENCE IN IRAQ DURING 2000-2015

Muzahem Mohammed AL-Hashimi* and Edrees M. Nori Mahmood

1Department of Statistics and Informatics, College of Computer science & Mathematics, University of Mosul, Mosul, Iraq
2Department of Operation Research and Intelligent Techniques, College of Computer science & Mathematics, University of Mosul, Mosul, Iraq

Corresponding author: Muzahem Mohammed AL-Hashimi
E-mail: muzahim_alhashime@uomosul.edu.iq, OR muzahim63@gmail.com

ABSTRACT

The spatial statistical analysis of breast cancer incidences across Iraq has not been explored in Iraq. This paper aimed to explore the spatial pattern and risk clusters of female breast cancer incidence from 2000 to 2015 in Iraq (except the Kurdish region). To enhance statistical stability and to access the changes over time, we split the data according to the geographical district into three periods (2000-2004, 2005-2009, and 2010-2015). The age-standardized incidence rates (ASRs) were calculated using the world standard population. Having obtained estimates ASRs, the global index of spatial autocorrelation (Moran’s I statistic) was used to assess spatial dependence across districts for ASRs. Anselin local Moran’s I statistic was used to identify spatial outliers. Additionally, we used Getis-Ord Gi* statistic to detect hotspots and coldspots over entire Iraq that represent clusters of districts with significantly high or low ASRs. A total of 44,496 cases were reported in 2000-2015 in Iraq, with an ASR of incidence of (32.81/100,000). The ASR incidence of breast cancer showed a significant average percentage change of 5.40% from 2002 to 2015. The spatial autocorrelation analysis showed insignificant positive spatial autocorrelation in 2000-2004, and significant positive spatial autocorrelation in 2005-2009, and in 2010-2015. This study identified four districts as high-risk areas for breast cancer during the two periods 2005-2009 and 2010-2015, including Al-Karkh, Al-Adhamia, Al-Rissafl, and Al-Sadr. This information can assist the allocation of health care resources and expand cancer prevention efforts.

Keywords: Hotspots, Coldspots, Spatial Outliers, Breast Cancer, Iraq

INTRODUCTION

Breast cancer in females is the most common cancer diagnosed worldwide, and the leading cause of cancer death. Globally, there were about 2.3 (11.7% of all sites) million newly diagnosed cases of breast cancer and about 0.7 (6.9% of all sites) million deaths in females in 20201. A 2020 GLOBOCAN report highlighted the large geographical diversity of major cancers observed in 185 countries1. The highest incidence rates were in Australia/New Zealand and Western Europe, and the lowest incidences were in middle Africa and south-central Asia, and the females in very high Human Development Index (HDI) countries have a higher breast cancer ASR (55.9/100,000) than females in either low HDI countries (29.7/100,000)1. The substantial variations in breast cancer rates may reflect changes in risk factors associated with economic development and availability of early detection and timely treatment2. According to the Iraqi Cancer Registry, breast cancer ranked first for both incidence and mortality. While the incidence of breast cancer among females in Iraq was relatively less than that in developed countries, we witnessed a substantially increasing incidence during the period 2002-2015.

The Iraqi Cancer Registry was established in 1974. It is located in the Institute of Radiology and Nuclear Medicine and supervised directly by the Iraqi Cancer Board. The information of every new cancer patient was collected from Baghdad province until 1989, then the registry services were extended to cover all provinces of Iraq. Cases registered were reported from inpatients and outpatients of government hospitals and cases attending the institute of Radiology and Nuclear Medicine. In addition, private hospitals, pathological and hematological laboratories requested to report to the registry all cases of cancer that come to their attention. The Iraqi Cancer Registry annual report provides information about cancer in Iraq including new cases of cancer

Iraq is a country in the southwest of Asia, which consists of 18 provinces according to Iraqi classification of administrative regions, three of which are located in the Kurdish region. It covers an area of 437,072 square kilometer\(^4\) and the population according to the World Population Review was over 39 million in 2019. Fig. 1 shows the location of Iraq.

Spatial analysis has been commonly used in health studies, such as epidemiology\(^5\). Spatial epidemiology is “the description and analysis of the geographical distribution of disease” \(^6\). Understanding the geographical distribution of disease in a population can provide important insight into the causes and controls of disease. Disease maps provide visual representations and an effective tool to show a huge amount of geographical information, it can be used to display patterns of diseases for defined geographical area\(^7\). This paper aims to explore the spatial distribution pattern and risk clusters of female breast cancer in different districts in Iraq during (2000-2015) (except the Kurdish region).

**MATERIALS AND METHODS**

The incidences of female breast cancer (code C50- based on the International Classification of Diseases for Oncology, Third Edition (ICDO-3) between 2000 and 2015 were obtained from the Iraqi Cancer Registry. Annual female population by 5-year age groups, gender, and provinces were obtained from the Ministry of Planning/Central Organization for Statistics. This data covers all provinces with 83 districts in the country, except three provinces in the Kurdish region (Erbil, Duhok, and Al-Sulaymaniyah, for which the data is incomplete).

To evaluate the incidence of female breast cancer during the study period 2000-2015, we split the data according to the geographical district into three periods (2000-2004, 2005-2009, and 2010-2015) to enhance statistical stability and to access the changes over time. The ASRs were calculated using the world standard population. Having obtained estimates ASR, two spatial statistical tools were applied to evaluate spatial clustering and outliers, the global spatial autocorrelation (i.e., Moran’s I) and local indicators of spatial association (i.e., Anselin local Moran’s I and Getis-Ord Gi\(^*\)). The descriptive analysis was carried out using SAS statistical software, version 9.4. As for visualizing the geographical district differences in the incidences, the maps of the breast cancer were created using ArcGIS software, version 10.6.

**Statistical methods**

*Global spatial autocorrelation - Moran’s I*

Moran’s global statistic (Moran’s I) was used to estimate the spatial autocorrelation. This statistic measures the similarity of the overall
area with respect to ASR. The global Moran’s I statistic is given as follows:

\[
I = \frac{n \sum_{i} \sum_{j} W_{ij} (X_i - \bar{X})(X_j - \bar{X})}{\sum_{i} \sum_{j} W_{ij} (X_i - \bar{X})^2}
\]

Where:
- \(X_i\) = the ASRs of breast cancer for the \(i\)th district;
- \(X_j\) = the ASRs of breast cancer for the \(j\)th district;
- \(\bar{X}\) = the mean of ASRs of breast cancer for all of the districts in the study area;
- \(W_{ij}\) = the spatial weight between pair of districts \(i\) and \(j\);
- \(n\) = the total number of districts.

The values of Moran’s \(I\) ranged from -1 to 1, where large and positive values indicate the presence of significant spatial autocorrelation \((p < 0.05)\). Low values of Moran’s \(I\) indicate the presence of significant regularity, while negative values suggest the clustering of dissimilar values. The global Moran’s \(I\) do not suggest where clusters of low or high ASR might appear. Thus, local tests for spatial autocorrelation such as Getis-Ord Gi* and Anselin Local Moran’s \(I\) are needed.

**Local indicators of spatial association**

We used two local measures of spatial association within ArcGIS 10.6 to examine the clusters of geographical districts with high or low ASRs or outliers are located and in addition, what type of spatial correlation is more important.

**The Anselin local Moran’s \(I\)**

The local measure of Moran’s \(I\) allowed us to identify spatial clusters of districts and spatial local outliers that are different from the neighboring districts. Based on the Moran’s \(I\) statistic and corresponding \((p – value)\), A high positive Moran’s \(I\) value (hotspot) refers to high ASR and surrounded by a high ASR (high-high), low positive Moran’s \(I\) value (coldspot) refers to low ASR and surrounded by a low ASR (low-low), A negative Moran’s \(I\) value (high-low) refers to high ASR and surrounded by a low ASR, low negative Moran’s \(I\) value (low-high) refers to low ASR and surrounded by a high ASR. The Anselin Local Moran’s \(I\) statistic is given as follows:

\[
I_i = \frac{(X_i - \bar{X}) \sum_j W_{ij} (X_j - \bar{X})}{S^2}
\]

Where:
- \(X_i\) = the ASRs of breast cancer for the \(i\)th district;
- \(\bar{X}\) = the mean of ASRs of breast cancer for all of the districts in the study area;
- \(W_{ij}\) = the spatial weight between pair of districts \(i\) and \(j\);
- \(n\) = the total number of districts; and
- \(S\) = the standard deviation of the ASRs of breast cancer in the study area.

**The Getis-Ord Gi* statistic**

Getis-Ord Gi* statistic was applied to examine spatial clustering. The Getis-Ord Gi* statistic is given as follows:

\[
G_i^* = \frac{\sum_{j=1}^{n} W_{ij} X_j - \bar{X} \sum_{j=1}^{n} W_{ij}}{S \sqrt{n \sum_{j=1}^{n} W_{ij}^2 - \sum_{j=1}^{n} W_{ij} (\bar{X}) - 1}}
\]

Where:
- \(X_i\) = the ASRs of breast cancer for the \(i\)th district;
- \(\bar{X}\) = the mean of ASRs of breast cancer for all of the districts in the study area;
- \(W_{ij}\) = the spatial weight between pair of districts \(i\) and \(j\);
- \(n\) = the total number of districts; and
- \(S\) = the standard deviation of the ASRs of breast cancer in the study area.

Getis-Ord statistic generates a z-value and associated \((p – value)\) for each district, where high \(z\)-value \((z > 1.96)\) indicates a significant hotspot and low \(z\)-value \((z < -1.96)\) indicate a significant coldspot \((p < 0.05)\).

**RESULTS**

**Basic information**

Between 2000 and 2015, 44496 female breast cancer cases were reported in Iraq, which accounted for 34% of the cancer cases among females, and it ranked as the first most common type of female cancer in Iraq. The number of newly diagnosed Iraqi female breast cancer increased from 1653 in 2000 to 4720 in 2015, which corresponds to a 2.86-fold
increase. ASR (all-age) for female breast cancer in Iraq for the period 2000 to 2015 was (32.81/100,000). The national ASR increased from (22.07/100,000) in 2000-2004 to (30.12/100,000) in 2005-2009 and to (41.07/100,000) female populations in 2010-2015 (Fig. 2).

**Fig. 2 Age-standardized incidence rate of breast cancer in Iraq over time**

The geographical distribution of breast cancer across provinces of Iraq is shown in (Fig 3). The ASR of breast cancer was 11.45-31.33 (31.33 per 100,000 females was derived in Baghdad) in 2000-2004 (Fig. 3A), 12.82-48.42 (48.42 per 100,000 females was derived in Al-Najaf) in 2005-2009 (Fig. 3B), and 23.90-55.27 (55.27 per 100,000 females was derived in Kirkuk) during the period 2010-2015 (Fig. 3C). Baghdad and Kirkuk provinces stand out with the highest ASR of breast cancer for the three study periods. The provinces Al-Najaf and Al-Basrah had the highest ASR for the period 2005-2009 along with Baghdad and Kirkuk provinces. The provinces Al-Anbar, Maysan, and Thi-Qar had the lowest ASR for the three study periods.

The geographical distribution of breast cancer across districts of Iraq is shown in Fig. 4. The ASR of breast cancer was 0.001-56.45 (56.45 per 100,000 females was derived in Al-Karkh district in Baghdad province) in 2000-2004 (Fig. 4A), 0.00-65.36 (65.36 per 100,000 females was derived in Al-Najaf city in Al-Najaf province) in 2005-2009 (Fig. 4B), and 1.50-99.37 (99.37 per 100,000 females was derived in Al-Karkh district in Baghdad province) in 2010-2015 (Fig. 4C).

**Temporal trend**

The Joinpoint analysis indicated that the ASRs had a significant increasing trend from 2000 to 2015, with an average percentage change of 6.10% (Fig. 5A). Furthermore, the trend (Fig. 5B) can be divided into two periods: the first period was from 2000 to 2002, during which the ASRs was increasing insignificantly with an APC of 17.14%, followed by a significant rapid growth period (2002-2015) with an APC of 5.40%.

The global Moran’s I test statistics and the associated z – value of the ASR in geographical districts were calculated for each of the three periods. The results across the time periods were different, with insignificant positive spatial autocorrelation (Moran’s I index = 0.112 (z = 1.886, P < 0.059)) for the period 2000-2004, and significant positive spatial autocorrelation (Moran’s I index = 0.211(z = 3.358, P < 0.008) for the period 2005-2009, and (Moran’s I index = 0.259 (z = 4.120, P < 0.001)) for the period 2010-2015. In other words, across Iraq, districts with similar ASR of

---

1 0.00 represent no cases were reported in some districts in Iraq during the study period.

Fig. 3 The geographical distribution of breast cancer incidence (per 100,000 females at risk) by provinces for the periods (2000-2004), (2005-2009), and (2010-2015). Age-standardized incidence rates for (2000-2004) (A), Age-standardized incidence rates for (2005-2009) (B), Age-standardized incidence rates for (2010-2015) (C)
Fig. 4 The geographical distribution of breast cancer incidence (per 100,000 females at risk) by districts for the periods (2000-2004), (2005-2009), and (2010-2015). Age-standardized incidence rates for (2000-2004) (A), Age-standardized incidence rates for (2005-2009) (B), Age-standardized incidence rates for (2010-2015) (C)
The results of Anselin’s Local Moran’s I (cluster and outlier analysis) for the period 2005-2009 suggest the presence of hotspots (high-high) of breast cancer across districts, Shat Al-Arab (Basrah province) in the south, Al-Hindiya (Karbala province), Al-Kufa (Al-Najaf province), and Al-Sadir, Al-Rissafa, Al-Karkh and Al-Adhamia (Baghdad province) in the middle of Iraq. Coldspots (low-low) were seen in Rawa (Al-Anbar province) and Sinjar (Nineva province) (Fig. 6A). These patterns were changed during the period 2010-2015, Shat Al-Arab and Al-Hindiya, which were no longer a hotspot. Additionally, hotspots districts appeared in Bakooba district. Coldspots were seen in Rawa (Al-Anbar province), Sinjar, Al-Hatra, and Al-Baaj (Nineva province) (Fig. 6B). In 2005-2009, one district was categorized as (low-high), Dibis (Kirkuk province) in the north of Iraq. In 2010-2015, the (low-high) were located in Al-Mada’in, Abu-Graib (Baghdad province), and Al-Hindiya (Kerbala province). In 2005-2009, there were three districts categorized as (high-low), included Al-Kut (Wassit province), Al-Nassiryah (Thi-Qar province), and Al-Samawa (Muthana province), where these districts showed high breast cancer incidence but surrounded by low incidence districts (Fig. 6A). This can be seen also over the period (2010-2015) except Al-Kut which no longer represents (high-low) type.

Additionally, the (high-low) type appeared in Ana district (Al-Anbar province) (Fig. 6B).

The results of the Getis-Ord Gi* can be seen in (Fig. 7A) and (Fig. 7B). During the period (2005-2009), 10 districts were identified as a hotspot. During the period (2010-2015), 4 districts were identified as a hotspot, and 4 districts were identified as a coldspot.

**DISCUSSION**

This paper is the first to describe province and district-level variation in breast cancer incidence rates over entire Iraq (except the Kurdish region), using spatial statistical tools to determine statistically significant spatial clusters of hotspots, coldspots, and spatial outliers in breast cancer incidence rates. According to the Iraqi Cancer Registry, breast cancer ranked first for both incidence and mortality. While the incidence of breast cancer among females in Iraq was relatively less than that in developed countries, we witnessed a substantially increasing incidence during the period 2010-2015. Results showed that the ASR of breast cancer was increased as a whole, from (15.40/100,000) in 2000 to (42.52/100,000) in 2015. There is geographical variation in Iraq breast cancer incidences according to provinces and district-level. This
study showed that the breast cancer by the province has increased from 2000-2004 to 2010-2015, there was a clear trend in cancer incidence. The highest ASR of breast cancer was observed in Baghdad and Kirkuk provinces in (2000-2004). In (2005-2009), the highest ASR of breast cancer was observed in the Al-Najaf, Al-Basrah, Baghdad, and Kirkuk provinces, while in (2010-2015), the highest ASR of breast cancer was observed in Baghdad and Kirkuk Provinces. The highest incidence was observed in the Al-Karkh district in Baghdad province during the period (2010-2015).

Fig. 6 Clusters and spatial outliers in breast cancer ASRs. clusters and spatial outliers in breast cancer ASRs, 2005 to 2009 (A), clusters and spatial outliers in breast cancer ASRs, 2010 to 2015 (B)
Fig. 7 Hotspots and coldspots for female breast cancer ASRs in Iraq. Hotspots and coldspots for the period 2005 to 2009 (A), Hotspots and coldspots for the period 2010 to 2015 (B)
A study in the Kurdistan region in Iraq (North of Iraq), reported that the (ASR) of breast cancer in Kurdish women was (17.9/100,000) during the period (2011-2013)\textsuperscript{11}, which is lower than the (ASR) for all provinces bordering the Kurdish area (Ninewa (32.21/100,000), Salah Al-Din (27.71/100,000), Diyala (31.81/100,000), and Kirkuk (55.27/100,000) during the period (2010-2015). The (ASR) of breast cancer in the central city of Sulaimaniyah women in the Kurdish region was (36.00/100,000) in 2012\textsuperscript{12}. This result higher than Salah Al-Din and Diyala provinces, but lower than the (ASR) in Kirkuk during the period (2010-2015).

A study in Basrah reported that the ASR of female breast cancer was 34.86 per 100,000 during the period (2009-2012)\textsuperscript{13}, which is lower than the ASR (38.1516 per 100,000 during the period (2005-2009), and 49.916 per 100,000 females during the period (2010-2015)) in the present study. Our finding is not consistent with the previous result. This inconsistency may be due to the difference in the study periods, or the difference between the methods of calculating (ASR). This paper adopted world standard population to adjust age effects on a geographical scale to calculate (ASR), while the previous study was not referred to the method of calculation.

The (ASR) in Iraq during the period (2010-2015) was higher than that in Saudi Arabia (22.1/100,000)\textsuperscript{14}, Oman (22.1/100,000)\textsuperscript{15}, Iran (29.83/100,000)\textsuperscript{16}. It less than that in Lebanon (96.5/100,000)\textsuperscript{17}, Jordan (45.7/100,000)\textsuperscript{18}, Bahrain (47.4/100,000)\textsuperscript{19}, Qatar (48.2/100,000)\textsuperscript{20} and Kuwait (46.7 /100,000)\textsuperscript{21}, but not lower than the ASR in Basrah Province (49.9160 /100,000 during the period 2000-2015), which is located in southern Iraq, bordering Kuwait. The (ASR) in Iraq was lower than the ASR in Korea (49.5/100,000)\textsuperscript{22}, Eastern Europe (57.1/100,000), Northern America (89.4/100,000), Western Asia (46.6/100,000), and Worldwide (47.8/100,000)\textsuperscript{1}.

The results of global Moran’s I indicated a spatial dependence of female breast cancer across Iraq districts. In other words, districts with high ASR tend to cluster together. The results of Getis Ord Gi\textsuperscript{+} identified the hotspots and coldspots across Iraq that revealed the existence of high-risk areas for breast cancer in Al-Karkh, Al-Adhamia, Al-Rissa, and Al-Sadir in Baghdad province (capital of Iraq), during the two periods 2005-2009 and 2010-2015. However, it is beyond the aim of this paper to know the reasons for the geographical variation of breast cancer incidence in Iraq. Recent studies indicate that exposure to environmental risk factors, such as certain chemicals during the development stages of breast tissue before birth and until menopause, may increase the risk of breast cancer later in life\textsuperscript{23}. The substantial variations in breast cancer rates may reflect changes in risk factors associated with economic development and availability of early detection and timely treatment\textsuperscript{2}. The variation in cancer incidence rates between the provinces in Iraq might be explained by environmental risk factors, such as chemicals exposures, and occupational exposures, especially most of the hotspots were detected in capital cities. Another reason for higher rates in some provinces might be that the oncology centers are located in these provinces and the cases occurring in the province were registered, while might be the cases in other provinces not reported. However, the reasons for the higher incidence rates in some districts merits further investigation.

CONCLUSION

In summary, there is a geographic variation in the ASR of breast cancer across Iraq. Districts with low and high breast cancer ASR tend to cluster together. This study identified 4 districts as high-risk areas for breast cancer during the two periods 2005-2009 and 2010-2015, including Al-Karkh, Al-Adhamia, Al-Rissafa, and Al-Sadir, and we have evidenced an increase in breast cancer incidence rates during 2000-2015. Measures targeting these high-risk areas should be developed and reduce geographical variation in the burden of breast cancer and allocation of health resources. More researches are needed to investigate the reasons for the geographic and temporal variations of breast cancer incidence in Iraq.

ACKNOWLEDGMENTS

We would like to express our thanks to the staff of the Iraqi Cancer Board-Ministry of Health for their effort in providing data access. We are
grateful to the University of Mosul, for supporting our work.

Conflict of interest
We declare that we have no conflicts of interest to disclose regarding this manuscript.

Funding
This manuscript received no specific grant from any funding agency in the public or commercial sectors.

REFERENCES


11. Karim SA, Ghalib HH, Mohammed SA, Fattah FH. The incidence, age at diagnosis of breast cancer in the Iraqi Kurdish population and comparison to some other countries of Middle-East and West.


23. [National Institute of Environmental Health Sciences. Breast cancer risk and environmental factors.]