Impact of sex, body mass index and initial pathologic diagnosis age on the incidence and prognosis of different types of cancer

XUAN HUANG[1*], CHUANJUN SHU[2*], LI CHEN[1*] and BING YAO[1]

[1]Reproductive Medical Center, Jinling Hospital Affiliated to The Medical School of Nanjing University, Nanjing, Jiangsu 210002; [2]State Key Laboratory of Bioelectronics, School of Biological Science and Medical Engineering, Southeast University, Nanjing, Jiangsu 210096, P.R. China

Abstract. Cancer represents a significant challenge for humankind, as early diagnosis and treatment are difficult to achieve. To systemically investigate the effect of sex, body mass index (BMI) and age on cancer incidence and prognosis, the data from 14,504 cases of cancer were downloaded from The Cancer Genome Atlas (TCGA). BMI was used to categorize each person as underweight, normal weight, overweight or obese. Two- and five-year survival rates were applied to estimate the prognosis for each cancer type. All data were statistically analyzed. We identified that males were more susceptible to lung, liver and skin cancer when compared with females, whereas females were more susceptible to thyroid, breast and adrenal cortex cancer. High BMI (>25) was positively associated with the occurrence of cancer, although patients with high BMI at the time of initial diagnosis had higher two/five-year survival rates. The survival rates for cancer were positively correlated with the age at initial pathologic diagnosis. Some types of cancer were associated with particularly young ages of onset, including adrenocortical carcinoma, cervical and endocervical cancers, brain lower grade glioma, pheochromocytoma and paraganglioma, testicular germ cell tumors and thyroid carcinoma. Hence, the early diagnosis and prognosis for these cancers need to be improved. In conclusion, sex, BMI and age are associated with the incidence and survival rates for cancers. These results could be used to supplement precision and personalized medicine.

Correspondence to: Professor Bing Yao, Reproductive Medical Center, Jinling Hospital Affiliated to The Medical School of Nanjing University, 305 Zhongshan Road, Nanjing, Jiangsu 210002, P.R. China
E-mail: yaobingnju@163.com

*Contributed equally

Abbreviations: ACC, adrenocortical carcinoma; BLCA, bladder urothelial carcinoma; BRCA, breast invasive carcinoma; CESC, cervical and endocervical cancers; CHOL, cholangiocarcinoma; COAD, colon adenocarcinoma; COADREAD, colorectal adenocarcinoma; DLBC, lymphoid neoplasm diffuse large B-cell lymphoma; ESCA, esophageal carcinoma; FPPP, FFPE Pilot Phase II; GBM, glioblastoma multiforme; GBMLGG, glioma; HNSC, head and neck squamous cell carcinoma; KICH, kidney chromophobe; KIPAN, pan-kidney cohort (KICH+KIRC+KIRP); KIRC, kidney renal clear cell carcinoma; KIRP, kidney renal papillary cell carcinoma; LAML, acute myeloid leukemia; LGG, brain lower grade glioma; LIHC, liver hepatocellular carcinoma; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; MESO, mesothelioma; OV, ovarian serous cystadenocarcinoma; PAAD, pancreatic adenocarcinoma; PCPG, pheochromocytoma and paraganglioma; PRAD, prostate adenocarcinoma; READ, rectum adenocarcinoma; SARC, sarcoma; SKCM, skin cutaneous melanoma; STAD, stomach adenocarcinoma; STES, stomach and esophageal carcinoma; TGCT, testicular germ cell tumors; THCA, thyroid carcinoma; THYM, thymoma; UCEC, uterine corpus endometrial carcinoma; UCS, uterine carcinosarcoma; UVM, uveal melanoma

Key words: cancer, sex disparity, body mass index, survival rates, initial pathologic diagnosis, ages

Introduction

Cancers are diseases involving the uncontrollable growth of abnormal cells that overcome the usual limitations to cell division. Cancer is now a relatively common disease. For example, there were about 90.5 million individuals diagnosed with cancer in 2015 (1), and more than 14.1 million new cases of cancer occur each year. Cancer is a leading cause of death worldwide, accounting for 8.8 million deaths in 2015; 15.7% of all deaths. The most common causes for cancer-related death are lung, liver, colorectal, stomach and breast cancers (2). Cancer is becoming an enormous burden on society in high- and low-income countries alike, costing an estimated 1.16 trillion USD worldwide in 2010 (3).

The global incidence of cancer is increasing due to the growth and aging of the population, as well as an increasing prevalence of established risk factors such as hormone replacement therapy, obesity, physical inactivity, radiation and chemical exposure, infection, autoimmune disease, smoking

and changing reproductive patterns, as are associated with urbanization and economic development (4). Some hormones play a role in the development of cancer by promoting cell proliferation. For example, a previous study indicated that there is a positive correlation between testosterone level and prostate cancer (5). Hormones are also important agents in other sex-specific cancers, such as cancers of the ovary, testis, endometrium and breast. Furthermore, sex disparity in the incidence, aggressiveness, response to therapy and prognosis has been observed for a variety of cancers, such as lung, prostate, colorectal, stomach, breast and uterine cancer (6).

As cancer is widespread and lethal, it is considered by some to be the most significant health problem faced by humanity. The majority of cancer deaths are associated with metastasis, but the mechanisms for metastasis and cancer spread are poorly characterized (7). A recent study indicated that dietary lipids may promote the metastasis of cancer cells (8). Adipocytes and their functionally related cells are strong candidates for the promotion of carcinogenesis and the influencing of tumor behavior (9). Body mass index (BMI) is used as a general measure of mass and can serve as an unspecific means of estimating adiposity (10). Obesity is a growing public health issue and the second most common preventable cause of death (11). It is estimated that over one billion adults are overweight and 315 million are obese worldwide (12). Economically developed countries have the highest prevalence of obesity. In USA, around one in three (36%) adults is obese (12). Researchers have identified eight types of cancer linked to excess weight and obesity: Stomach, liver, gall bladder, pancreas, ovary, thyroid cancer, meningioma and multiple myeloma (13). However, to the best of our knowledge, the relationship between BMI and cancer has not been systemically investigated.

Cancer survival rates, including two- and five-year survival, are used to estimate the prognosis for a specific cancer (14,15). The chance of survival depends on the type of cancer and the extent of disease at the start of treatment (16). The early detection of cancer is associated with a good prognosis (16,17). Recently, with developments in the treatment of cancer, the survival rates for patients with various types of cancer have been improved (18-21). For example, the overall mortality rate for cancer in the USA has dropped by at least one-fifth over the past two decades according to new statistics from the American Cancer Society (22). The improvements in survival rates can also be attributed to improvements in diagnosis and prognosis (16,23). The importance of patient characteristics to long-term survival is the subject of much debate (24-26). The factors affecting cancer survival rates include sex, age, geographical location and the pathological type of cancer (27-30). The most significant risk factor for developing cancer is age. The effect of aging on cancer is complicated by factors such as DNA damage and inflammation promoting tumorigenesis, and factors such as vascular aging and endocrine changes inhibiting it (31).

The epidemiology of cancer has changed with the advancement of technology and social development (32). By using The Cancer Genome Atlas database (TCGA), the present study was designed to systemically investigate the effect of sex, BMI and age on cancer incidence and survival rates for various cancer types. It is anticipated that clarifying these issues will lead to a deeper understanding of cancer, as will be necessary for the development of precision and personalized medicine.

Materials and methods

Clinical data preparation. TCGA is a project, started in 2005, to improve the diagnosis, treatment and prevention of cancer through a better understanding of its genetic basis. This project is supervised by the National Cancer Institute's Center for Cancer Genomics and the National Human Genome Research Institute, which are funded by the US government. Clinical data of the patients from the TCGA database (version 2016_01_28, https://cancergenome.nih.gov/) were downloaded using the Broad GDAC Firehose (http://gdac.broadinstitute.org/). We downloaded the clinical data of 14,504 cancer cases, including 38 different types of cancer. The clinical data included patient ID number, vital status, overall survival time, age at initial pathologic diagnosis, time to last follow-up, gender, height and weight.

Body mass index (BMI). BMI, a measure calculated from height and weight, was used to categorize a person as underweight, normal weight, overweight, or obese (33). We included only BMI measurements taken after the age of 20, assuming that subjects had reached their adult height by this age. BMI was computed as each patient's weight in kilograms divided by the square of their height in meters (kg/m^2). The patients were principally divided into 3 groups on the basis of BMI: Underweight (<18.5 kg/m^2), normal (≥18.5 to <25.0 kg/m^2) and high (≥25.0 kg/m^2) (33). The high group included two sub-groups: Overweight (≥25 to <30 kg/m^2) and obese (≥30 kg/m^2).

Two- and five-year survival rates. The survival rate is the percentage of individuals with a disease still alive after a given period of time after the initial diagnosis or treatment, which can be used to estimate the prognosis for that disease (27). An overall survival rate for a type of cancer includes individuals of all ages and health conditions, including those diagnosed particularly early or late. There are various types of survival rates, but two types of survival rates, i.e. two- and five-year survival rates, are more commonly cited in cancer statistics (34,35). For aggressive cancers with a shorter life expectancy following diagnosis, two-year survival statistics are typically used to estimate prognosis. Some cancers can recur many years after their initial identification and treatment. If they have not recurred by five years after the initial pathogenic diagnosis, the chance of a later recurrence is very small (36). Therefore, the five-year survival rate can be used to compare the effectiveness of treatment. In the present study, two- and five-year survival rates for each of 38 types of cancers were calculated by Kaplan Meier curves and were compared by performing the Tarone-Ware test.

Statistical analysis. Statistical analysis system (SAS) is a software suite that can mine, alter, manage and retrieve data from a variety of sources and perform statistical analysis (37). The SAS software suite has >200 components. The SAS/STAT component was used to perform the statistical analysis in this present study. A t-test was used to determine if two sets of
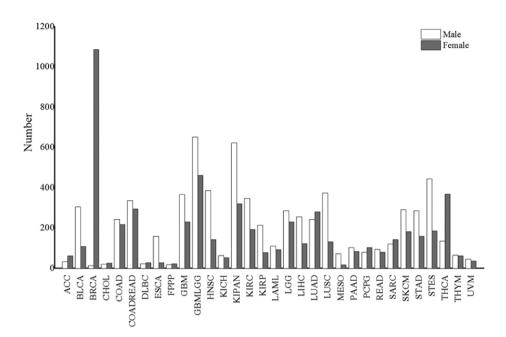
Figure 1. The number of cases for different types of cancer. The white and gray bars represent males and females, respectively.

data were significantly different from each other, such as sex and BMI for different types of cancer (38). The chi-squared test was used to determine whether there was a significant difference between the expected and observed frequencies in one or more categories, such as incidence, two- and five-year survival rate for different types of cancer (39). The P-value is the probability for a given statistical model that, when the null hypothesis is true, the statistical summary (such as the sample mean difference between two compared groups) would be the same as or of greater magnitude than the actual observed results. The cut-off for statistical significance was set as P<0.05. We performed a multiple testing correction procedure (Bonferroni adjustment) to adjust our statistical confidence intervals based on the number of tests performed. Origin is a computer program for interactive scientific graphing and data analysis. The bar charts and line charts were produced with Origin 9.0 software (https://www.originlab.com). R 3.5.0 (https://www.r-project.org/) is a software for statistical computing and graphics (40). The pheatmap package of R (https://www.r-project.org/) was used to produce heat maps.

Results

Sex disparity is identified in the proportion of cancer incidence. There were 38 types of cancer and 14,504 cancer cases in TCGA version 2016_01_28. A sex disparity in cancer incidence, aggressiveness and prognosis has previously been observed in a variety of cancer types. To systemically investigate the sex disparity in cancer incidence, the number of cases for each of the 38 types of cancers were analyzed. In this TCGA version, 2 types of cancer were specific to males, including prostate adenocarcinoma (PRAD) and testicular germ cell tumors (TGCT) and 4 were specific to females, including cervical and endocervical cancers (CESC), uterine corpus endometrial carcinoma (UCEC), uterine carcinosarcoma (UCS) and ovarian serous cystadenocarcinoma (OV).

The sex ratios (male/female) for the incidence of all 38 cancers and the 32 non-sex-specific cancers were 1.04 and 1.2, respectively. This indicated that men are more susceptible to cancer compared with women. However, the result was reversed for some cancers, including breast invasive carcinoma (BRCA), thyroid carcinoma (THCA), adrenocortical carcinoma (ACC), lung adenocarcinoma (LUAD), pheochromocytoma and paraganglioma (PCPG) and sarcoma (SARC) (all P-values <0.024; P<0.05). In addition, there was no sex disparity in the proportion of cholangiocarcinoma (CHOL), lymphoid neoplasm diffuse large B-cell lymphoma (DLBC), kidney chromophobe (KICH), acute myeloid leukemia (LAML), pancreatic adenocarcinoma (PAAD), rectum adenocarcinoma (READ), thymoma (THYM) and uveal melanoma (UVM) (Fig. 1). The top 10 cancers types in males were glioma (GBMLGG), pan-kidney cohort (KIPAN), PRAD, stomach and esophageal carcinoma (STES), head and neck squamous cell carcinoma (HNSC), lung squamous cell carcinoma (LUSC), glioblastoma multiforme (GBM), kidney renal clear cell carcinoma (KIRC), colorectal adenocarcinoma (COADREAD) and bladder urothelial carcinoma (BLCA). Furthermore, the top 10 cancer types in females were BRCA, OV, UCEC, GBMLGG, THCA, KIPAN, CESC, COADREAD, LUAD and GBM (Fig. 1). This result indicates that four types of cancers, i.e. GBMLGG, KIPAN, GBM and COADREAD, are prevalent in both males and females.

To investigate the sex disparity in the cancer predilection site, a human body map for cancer was produced. As shown in Fig. 2, different colors were used to show the sex differences in the distribution of cancer in human body. According to this data version, an interesting phenomenon was identified, in that primary cancer did not occur in four organs, i.e. the heart, spleen, small intestine and epityphlon. There was also no primary cancer incidence in the precava, postcava, aorta and diaphragm muscle (Fig. 2). These organs are not typically associated with chronic inflammation (41). Hence, chronic
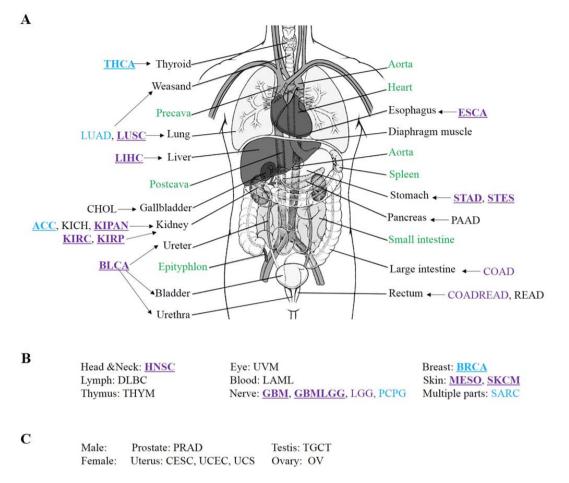
A

B

Head &Neck: **HNSC**	Eye: UVM	Breast: **BRCA**
Lymph: DLBC	Blood: LAML	Skin: **MESO**, **SKCM**
Thymus: THYM	Nerve: **GBM**, **GBMLGG**, LGG, PCPG	Multiple parts: SARC

C

Male:	Prostate: PRAD	Testis: TGCT
Female:	Uterus: CESC, UCEC, UCS	Ovary: OV

Figure 2. Sex differences in the distribution of human cancer. (A) Distribution of cancers in internal organs. (B) Distribution of cancers which do not occur in internal organs. (C) Male- and female-specific cancers. Purple and blue represent male and female, respectively. Underlining indicate a type of cancer with a significant sex disparity. Green indicates that no primary cancer data were available for these internal organs.

inflammation may be the most important cause of cancer, which is consistent with the results of previous studies (41,42). Males were more susceptible to cancers in the lung, liver, kidney, ureter, bladder, urethra, esophagus, stomach, head and neck, nervous system and skin, when compared with females. However, females were more susceptible to cancers in the thyroid, breast and adrenal cortex, when compared with males. These results indicated that predilection sites of cancer were different for males and females.

High BMI is associated with an increased risk of cancer. In this version of TCGA, there were 3,432 patients >20 years with recorded values for weight and height, from 18 types of cancer. According to the classification standard of the World Health Organization (WHO), for all 18 types of cancer, 64.42% of patients had a high BMI (≥25), including 32.34% overweight patients (BMI ≥25 to <30) and 32.08% obese patients (BMI ≥30). The mean percentage of patients with high BMI for each cancer type was a minimum of 56%, except in DLBC, esophageal carcinoma (ESCA), liver hepatocellular carcinoma (LIHC) and STES (Fig. 3). Furthermore, the mean percentage of patients with low BMI (BMI <18.5) for all types of cancer was 2.7% and did not exceed 7% for any type of cancer (Fig. 3). These results showed that high BMI (≥25) was likely to be a risk factor for many types of cancer (P=9.97E-28, P<0.05), with the exception of DLBC, ESCA, LIHC and STES
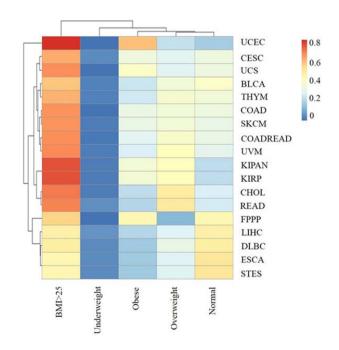

Figure 3. The association of BMI range with each type of cancer. The rows and columns represent the types of cancer and the BMI ranges. The labels represent the distribution of BMI range for each type of cancer (0-0.8162). BMI: underweight (<18.5 kg/m^2), normal (≥18.5 to <25.0 kg/m^2) and high (≥25.0 kg/m^2). The high group included two sub-groups: overweight (≥25 to <30 kg/m^2) and obese (≥30 kg/m^2).
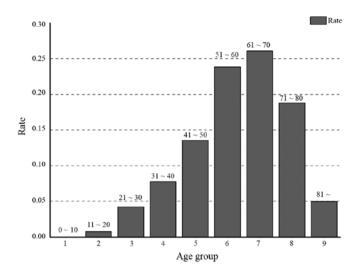
Figure 4. The distribution of age groups for all types of cancer. The age groups 1, 2, 3, 4, 5, 6, 7, 8 and 9 represent an initial pathologic diagnosis age of 0-10, 11-20, 21-30, 31-40, 41-50, 51-60, 61-70, 71-80 and >80, respectively.

and these 18 types of cancer were less likely in people with low BMI (<18.5).

As displayed in Fig. 3, patients with high BMI accounted for a particularly high proportion in CHOL, READ, KIPAN, kidney renal papillary cell carcinoma (KIRP) and were most predominant in UCES (81.6%). The result suggested that high BMI (≥25) was a high risk factor for cancer in the gallbladder, rectum, kidney and uterus. Furthermore, the proportion of obese patients with UCES and UCS were 59.6 and 38.5%. These results indicated that high BMI (≥30) was strongly associated with uterine cancer and that BMI affects different types of cancer in different ways.

The incidence of cancer demonstrates a young age trend. The majority of entries in the TCGA database have a record of the initial pathologic diagnosis age. The incidence of cancer in young adults has been increasing over the last 50 years (32). Early diagnosis is beneficial for the treatment and prognosis of cancer. To investigate which types of cancer are associated with younger patients, all cases in the TCGA database were grouped by the initial pathologic diagnosis age. They were divided into 9 groups: Group 1 (≤10) (no >10 years old); group 2 (10-20); group 3 (21-30); group 4 (31-40); group 5 (41-50); group 6 (51-60); group 7 (61-70); group 8 (71-80); and group 9 (>80 years). Cancer occurred in every age group. According to TCGA data, 0.01% of cancer cases occurred in patients ≤10 years old, and 4.9% of cases occurred in patients ≤30 years old (Fig. 4). With an increase in age, the relative distribution of cancer increased from group to group, with a peak in group 7. After the age of 70, the proportion of the cases of cancer decreased, likely due to the smaller proportion of the population at this age. After the age of 80, the proportion of cases of cancer decreased markedly, which can be attributed to the average life expectancy (43).

There were 2, 16 and 27 types of cancers in groups 1, 2 and 3, respectively (Fig. 5), indicating that many types of cancer occur in young people. These cancers occurred in sites throughout the body, including the kidney, liver, breast, stomach, uterus, gallbladder, lymph, thymus, esophagus, nerve, head and neck, ovary, blood, skin, testis and eye. In addition, 6 types of cancer were skewed towards occurring in young individuals: ACC, CESC, brain lower grade glioma (LGG), PCPG, TGCT and THCA. ACC, CESC, LGG, PCPG, TGCT and THCA were most common group 3 (16.3%), 4 (21.82%), 3 and 4 (17.09 and 30.29%), 3 and 4 (12.85 and 18.99%), 3 and 4 (40.30 and 37.31%), and group 3 and 4 (11.53 and 23.26%), respectively. In particular, 77.61% of TGCT patients were aged between 20 and 40. The results suggested that TGCT predominately occurs at reproductive age, which is potentially a significant finding for individuals of this age.

Sex, BMI and initial pathologic diagnosis age impact the survival rate in cancer. To investigate the factors that affect the survival rate, two- and five-year survival rates stratified by sex, BMI ranges, initial pathologic diagnosis age groups and cancer types were analyzed. We found that the two-year and five-year survival rates were different for each type of cancer (Fig. 6A). GBM and PAAD had the lowest two-year survival rates, at ≤30%. In addition, the two-year survival rates for LAML and mesothelioma (MESO) were in the range of 30-40%, and two-year survival rates for ESCA, STES, GBMLGG, stomach adenocarcinoma (STAD), BLCA and UCS were in the range of 40-50% (Fig. 6A). Twenty-two types of cancer were associated with a high (>70%) two-year survival rate, i.e. ACC, BRCA, CESC, colon adenocarcinoma (COAD), COADREAD, DLBC, KICH, KIPAN, KIRC, KIRP, LGG, OV, PCPG, PRAD, READ, SARC, skin cutaneous melanoma (SKCM), TGCT, THCA, THYM, UCEC and UVM. The 10 types of cancers with the lowest two-year survival rates occurred in the nervous system, pancreas, blood, skin, esophagus, stomach, ureter and uterus, which develop rapidly (such as PAAD and LAML), have previously been associated with a poor prognosis (such as MESO), or are not easy to find in early stages (such as GBM and GBMLGG) (16,23).

For the five-year survival rate, survival rates of GBM were ≤5%, rates for ESCA, MESO, LAML, PAAD, STES, CHOL, and STAD were in the 5-10% range, rates for UCS, UVM, and GBMLGG were in the 10-20% range. However, the five-year survival rate was ≥70% for BRCA, KICH, PCPG, PRAD, TGCT, THCA and THYM. These results indicated there are insufficient treatment options for GBM, ESCA, MESO, LAML, PAAD, STES, CHOL, STAD, UCS, UVM and GBMLGG. The five-year survival rate was lower than the two-year survival rate for all cancer types (Fig. 6B). The differences between the two- and five-year survival rates for each of CHOL, COAD, COADREAD, LGG, LUAD, OV, READ and UVM were ≥40%, indicating that these types of cancer have a lack of effective treatment available. The results suggested that the prognoses for these eight cancers would benefit from an early diagnosis process, as well as the development of new drugs to treat them. However, the difference for GBM, KICH, PCPG, PRAD, TGCT and THYM was in the 5-18% range. Furthermore, there was no statistical discrepancy between the two- and five-year survival rates for TGCT. The small differences for these types were attributed to two separate reasons: Firstly, the majority of cancer patients died within two years after initial pathologic diagnosis in some types, such as GBM. Secondly, other types of cancer in the list have an effective
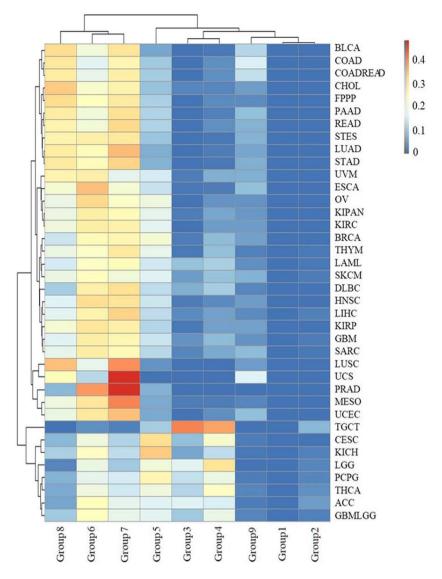
Figure 5. The distribution of age groups for each type of cancer. The rows and columns represent types of cancer and age groups. The color represents the proportion of the cases distributed in the age groups (0 to 0.4749). The age groups 1, 2, 3, 4, 5, 6, 7, 8 and 9 represent an initial pathologic diagnosis age of 0-10, 11-20, 21-30, 31-40, 41-50, 51-60, 61-70, 71-80 and >80, respectively.

treatment and prognosis, such as KICH, PCPG, PRAD, TGCT, THYM and particularly, TGCT.

It is necessary to identify which cancers have a large sex-specific difference, since previous studies have shown potential sex-specific differences in the pathophysiology, clinical presentation and treatment of cancers (44). The two- and five-year survival rates for males and females were different for all cancer types (Fig. 6C and D). Compared with males, females had a higher two-year survival rate for nine cancer types, i.e. BRCA, CHOL, ESCA, GBM, GBMLGG, LUAD, LUSC, STES and UVM, but lower two-year survival rates for three cancers, i.e. HNSC, KIRP and SARC (Fig. 6E). For other cancers, there was no significant difference between the two-year survival rate for females and males (for all other cancer types, P=0.19, P>0.05). For the five-year survival rate, the sex distribution changed. Compared with males, females had a higher five-year survival rate for COAD, COADREAD, ESCA, LGG, LUSC, MESO, PCPG, READ and THCA, but a lower five-year survival rate for DLBC, KIRP, LUAD and

THYM (Fig. 6E). For other cancer types, there was no significant difference in five-year survival rate between females and males (for all other cancer types, P=0.17, P>0.05).

In recent years, researchers have stressed the importance of maintaining a healthy BMI (10,33). Therefore, the distribution of survival rates for different BMI ranges were analyzed for each type of cancer. We identified a positive correlation between a high BMI and the corresponding two/five-year survival rate in cancer samples (r=0.53, Spearman correlation coefficient, Figs. 3 and 6A). In addition, as displayed in Fig. 6F, there was an interesting phenomenon in that patients with a high BMI (>25) had a higher two/five-year survival rate compared with patients with a normal BMI (18.5 to 25). The results may indicate that human with high BMI, who have more energy, could better tolerate a long process of physical deterioration (42).

The cancer survival rate was found to be related to the initial pathologic diagnosis age, i.e. the younger the pathologic diagnosis age, the better prognosis and the smaller mortality
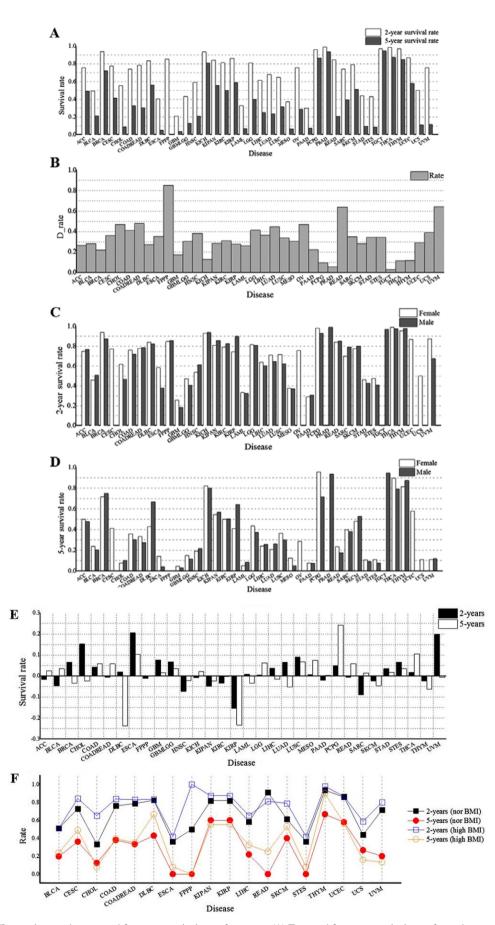
Figure 6. Sex and BMI range impact the two- and five-year survival rates for cancer. (A) Two- and five-year survival rates for each type of cancer. (B) The difference between two- and five-year survival rates for each type of cancer. The values of the y-axis represent the difference between two-year and corresponding five-year survival rates. (C and D) Two- and five-year survival rates for cancers in different sexes. (E) The difference in survival rate between males and females in each type of cancer. The values of the y-axis represent the survival rates for females minus the corresponding survival rates for males. (F) Survival rates for different ranges of BMI. Normal (Nor) BMI: 18.5-25 kg/cm^2, high BMI: >25 kg/cm^2. Different color lines represent different survival rates.
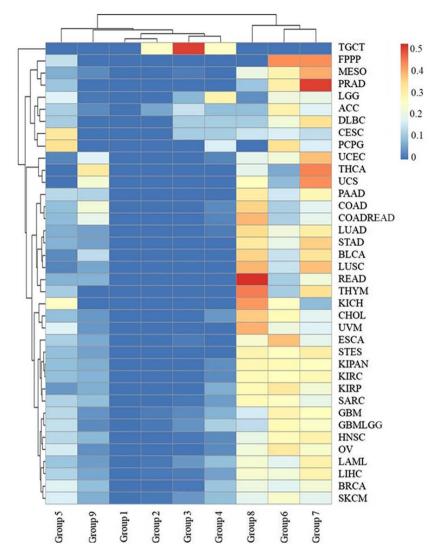
Figure 7. The distribution of the initial pathologic diagnosis age groups for the deceased patients with each type of cancer. The rows and columns represent the types of cancer and age groups. The color represents the proportion of the cases distributed in the age group (0-0.5185). The age groups 1, 2, 3, 4, 5, 6, 7, 8 and 9 represent an initial pathologic diagnosis age of 0-10, 11-20, 21-30, 31-40, 41-50, 51-60, 61-70, 71-80 and >80, respectively.

for all types of cancer with the exception of TGCT (Fig. 7). The age of patient death for TGCT was ≤40 years of age in all cases. Furthermore, >40% of deceased patients with LGG, CESC, PCPG and ACC were in age groups 1-5, and 30% deceased patients with DLBC, SKCM and GBMLGG were <50 years of age. These results indicated that more attention should be paid to the biochemical indicators of nerve, kidney, uterus, lymph and skin tumors in younger individuals. However, the initial pathologic diagnosis ages for deceased patients with most types of cancer were distributed in age groups 6, 7 and 8 (Fig. 7), including >80% for LUSC, PRAD, MESO, THYM, STAD, LUAD, READ, STES, BLCA, CHOL, KIRP, KIRC, KIPAN and ESCA.

Discussion

Previous studies have indicated a sex disparity in cancer incidence, aggressiveness, response to therapy and prognosis in a variety of cancers, including breast, endometrial, prostate, ovarian and testicular cancer (45,46). In the present study, we systemically observed the cancer gender disparity for 38 types

of cancer. Cancer sex disparity was not only identified for sex-related cancers, including cancers of the prostate, testis, uterus, ovary and breast, but also cancers of other organs, including the lung, liver, kidney, digestive tract, urethra, nerves, skin and head. The causes for the disparity may be the distinctive sex hormone expression between males and females, which plays an important role in cancer cell proliferation, apoptosis and differentiation, indicating a likely involvement in carcinogenesis. For example, high levels of testosterone, and estrogen and progesterone are risk factors for prostate cancer and breast cancer, respectively (47,48). The reasons of the disparity for other cancer types are not fully understood, and are likely due to lifestyle, psychological distress (anxiety, depression), social distress (family problems, job-related problems) and spiritual distress (spiritual pain, spiritual alienation) (46). For example, compared with females, males are more susceptible to cancers in the lung and liver, likely due to their increased smoking and alcohol consumption rates (49). These results indicated that it is necessary to differentiate cancer screening practices and treatments between females and males. During regular physical examinations, we should pay more attention to the

KIPAN, PRAD, STES, HNSC, LUSC, KIRC and BLCA for men, and the BRCA, OV, UCEC, THCA, KIPAN, CESC and LUAD for women.

The adult BMI of many countries is trending higher. A previous study indicated that the global obesity prevalence will reach 18% for men and 21% for women. Severe obesity will surpass 6% for men and 9% for women by 2025 (50). Being underweight, and severe and morbid obesity, are associated with elevated risks of adverse health outcomes (51). In the present study, we found that high BMI ($\leqq25$) is potentially a risk factor for many types of cancer. However, there were four types of cancer, i.e., DLBC, ESCA, LIHC and STES, which were associated with a low average BMI, which may be because the initial clinical symptoms for these cancers included weight loss, resulting in a low weight at initial pathologic diagnosis. Patients with a low BMI (<18.5) had a reduced incidence for all 38 types of cancer. These results indicated that adipocytes and their functionally related cells likely play important roles in the promotion of carcinogenesis. For example, adipocytes are thought to interact with cancer cells in the breast to promote the malignant growth of breast tumors (52). Furthermore, more and more research has confirmed a relationship between lipid metabolism and the proliferation of cancers (53). However, the mechanism of adipocytes and their functionally related cells in the promotion of carcinogenesis is not fully established and will be the core of our future work.

There is an increasing trend in young people in many types of cancer. For example, there was a large increase in breast cancer among young women in Brazil between 1988 and 2003 (54). The reason for the increase is uncertain. In the present study, we systemically investigated the incidence in young people in 38 types of cancer. A percentage of 4.9% of all cancer cases occurred in patients ≤30 years old. ACC, CESC, LGG, PCPG, TGCT and THCA were particularly associated with young patients. A total of 77.61% of the cases of TGCT occurred in patients that were 20-40 years old. ACC has a bimodal age distribution by age which is clustered in children under 5 years old and adults aged 30-40 years. CESC may be identified earlier due to the universal application of cervical cytology screening (55). LGG is the most common childhood brain tumor (56). PCPG and TCGT mostly occur in young or middle-age adults and TGCT most commonly occurs in between the ages of 35 and 65.

Survival for all types of cancer has markedly improved since the last century (57). A number of improvements in treatment during this period have undoubtedly contributed to the improved survival rate. Previous studies have shown sex, BMI and initial pathologic diagnosis age differences in the pathophysiology, clinical presentation and treatment response of cancers, such as head and neck cancer, endometrial cancer and breast cancer (10,44,58). In the present study, we systemically investigated the effect of sex, BMI and initial pathologic diagnosis age on two- and five-year survival rates for each of the 38 different types of cancer in TCGA. We first identified that two- and five-year survival rates were different for different types of cancer. Nine types of cancer, i.e., GBM, PAAD, LAML, MESO, ESCA, STES, GBMLGG, STAD and UCS, were associated with particularly low two- and five-year survival rates, indicating that the treatment of these types of

cancer should be improved. In addition, females generally had a higher survival rate than males, with the exception of two-year survival rates for HNSC, KIRP and SARC, and five-year survival rates for DLBC, KIRP, LUAD and THYM. Survival rates for these specific cancers vary widely between men and women, potentially due to biologically and socio-culturally determined factors. The role of sex and gender in cancer etiology, prevention and treatment is determined by complex, interacting variables that differ by cancer site. A comprehensive understanding requires the consideration of physiology, anatomy, hormones, behavior, lifestyle, environment and access to medical care.

Furthermore, there was a positive relationship between high BMI and survival rate. High BMI association with cancer is conflicting, as high BMI is a risk factor for many types of cancer. However, this result is consistent with the previous hypothesis that adipocytes not only play a critical role in the promotion of carcinogenesis, but that leukemia cell proliferation is inhibited when adipocytes in the bone marrow increase (59). This conflict may arise from adipocytes providing sufficient energy for normal and abnormal metabolic activity during cancer.

Regarding the initial pathologic diagnosis age, it is known that an earlier pathologic diagnosis of a cancer is associated with a better prognosis (60). The initial pathologic diagnosis age of the deceased cancer patients in the TCGA dataset was predominately distributed between the ages of 60 and 80. These results suggest that there is an increased chance of survival for most types of cancer when they are identified earlier. Cancers are associated with relatively low mortality when the initial pathologic diagnosis age is below 40, with the exception of TGCT. Furthermore, the initial pathologic diagnosis age for >30% of deceased patients in seven cancers, i.e., LGG, CESC, PCPG, ACC, DLBC, SKCM and GBMLGG, was below 50. These types of cancer were consistent with the types of cancers that were associated with a young age of diagnosis. For these types of cancers, we should strengthen early detection and rapid response systems to improve their prognosis.

The results of this study indicated that many factors impact the prognosis of cancers, such as sex, BMI and initial pathologic diagnosis age. Different types of cancer, sex and age groups may require different treatments. These results provide new insights, which may inform the development of precision and personalized medicine.

Availability of data and materials

The datasets used during the present study are available from the corresponding author upon reasonable request.

Authors' contributions

XH and CS designed and carried out the study and drafted the manuscript. LC designed the study and helped write the manuscript. BY and LC conceived the study and were the lead writers of the manuscript. All authors have read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. GBD 2015 Disease and Injury Incidence and Prevalence Collaborators: Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: A systematic analysis for the Global Burden of Disease Study 2015. Lancet 388: 1545-1602, 2016.
2. GBD 2015 Mortality and Causes of Death Collaborators: Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: A systematic analysis for the Global Burden of Disease Study 2015. Lancet 388: 1459-1544, 2016.
3. Enkhtaivan G, Kim DH and Pandurangan M: Cytotoxic effect of TDZ on human cervical cancer cells. J Photochem Photobiol B 173: 493-498, 2017.
4. Sud A, Kinnersley B and Houlston RS: Genome-wide association studies of cancer: Current insights and future perspectives. Nat Rev Cancer 17: 692-704, 2017.
5. Herati AS, Kohn TP, Butler PR and Lipshultz LI: Effects of testosterone on benign and malignant conditions of the prostate. Curr Sex Health Rep 9: 65-73, 2017.
6. Edgren G, Liang L, Adami HO and Chang ET: Enigmatic sex disparities in cancer incidence. Eur J Epidemiol 27: 187-196, 2012.
7. Micalizzi DS, Maheswaran S and Haber DA: A conduit to metastasis: Circulating tumor cell biology. Genes Dev 31: 1827-1840, 2017.
8. Pascual G, Avgustinova A, Mejetta S, Martin M, Castellanos A, Attolini CS, Berenguer A, Prats N, Toll A, Hueto JA, *et al*: Targeting metastasis-initiating cells through the fatty acid receptor CD36. Nature 541: 41-45, 2017.
9. Blee AM and Huang H: Fat lure: Adipocytes attract cancer cells out of the prostate. Translational Cancer Res 5 (Suppl 1): S123-S125, 2016.
10. Rojas KE, Matthews N, Raker C, Clark MA, Onstad M, Stuckey A and Gass J: Body mass index (BMI), postoperative appearance satisfaction, and sexual function in breast cancer survivorship. J Cancer Surviv 12: 127-133, 2018.
11. Amadou A, Ferrari P, Muwonge R, Moskal A, Biessy C, Romieu I and Hainaut P: Overweight, obesity and risk of premenopausal breast cancer according to ethnicity: A systematic review and dose-response meta-analysis. Obes Rev 14: 665-678, 2013.
12. Finucane MM, Stevens GA, Cowan MJ, Danaei G, Lin JK, Paciorek CJ, Singh GM, Gutierrez HR, Lu Y, Bahalim AN, *et al*: National, regional, and global trends in body-mass index since 1980: Systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9.1 million participants. Lancet 377: 557-567, 2011.
13. Rafee S, Flavin A and O'Reilly S: Body fatness and cancer. N Engl J Med 375: 2008, 2016.
14. Montazeri A: Quality of life data as prognostic indicators of survival in cancer patients: An overview of the literature from 1982 to 2008. Health Qual Life Outcomes 7: 102, 2009.
15. Moghimi-Dehkordi B and Safaee A: An overview of colorectal cancer survival rates and prognosis in Asia. World J Gastrointest Oncol 4: 71-75, 2012.
16. Tang Y, Qiao G, Xu E, Xuan Y, Liao M and Yin G: Biomarkers for early diagnosis, prognosis, prediction, and recurrence monitoring of non-small cell lung cancer. Onco Targets Ther 10: 4527-4534, 2017.
17. Gesthalter YB, Vick J, Steiling K and Spira A: Translating the transcriptome into tools for the early detection and prevention of lung cancer. Thorax 70: 476-481, 2015.
18. Li M, Wang MM, Guo XW, Wu CY, Li DR, Zhang X and Zhang PT: Different survival benefits of Chinese medicine for pancreatic cancer: How to choose? Chin J Integr Med 24: 178-184, 2018.
19. Wills B, Cardona AF, Rojas L, Ruiz-Patino A, Arrieta O, Reguart N, Carranza H, Vargas C, Otero J, Corrales L, *et al*: Survival outcomes according to TIMP1 and EGFR expression in heavily treated patients with advanced non-small cell lung cancer who received biweekly irinotecan plus bevacizumab. Anticancer Res 37: 6429-6436, 2017.
20. Mercier J and Voutsadakis IA: A Systematic review and Meta-analysis of retrospective series of regorafenib for treatment of metastatic colorectal cancer. Anticancer Res 37: 5925-5934, 2017.
21. Sineshaw HM, Jemal A, Thomas CJ and Mitin T: Changes in treatment patterns for patients with locally advanced rectal cancer in the United States over the past decade: An analysis from the National Cancer Data Base. Cancer 122: 1996-2003, 2016.
22. Siegel RL, Miller KD and Jemal A: Cancer statistics, 2017. CA Cancer J Clin 67: 7-30, 2017.
23. Shiani A, Narayanan S, Pena L and Friedman M: The role of diagnosis and treatment of underlying liver disease for the prognosis of primary liver cancer. Cancer Control 24, 2017.
24. Chennubhotla C, Clarke LP, Fedorov A, Foran D, Harris G, Helton E, Nordstrom R, Prior F, Rubin D, Saltz JH, *et al*: An assessment of imaging informatics for precision medicine in cancer. Yearb Med Inform 26: 110-119, 2017.
25. Brunner M, Olschewski M, Geibel A, Bode C and Zehender M: Long-term survival after pacemaker implantation. Prognostic importance of gender and baseline patient characteristics. Eur Heart J 25: 88-95, 2004.
26. Burney IA and Lakhtakia R: Precision medicine: Where have we reached and where are we headed? Sultan Qaboos Univ Med J 17: e255-e258, 2017.
27. Yavas O, Hayran M and Ozisik Y: Factors affecting survival in breast cancer patients following bone metastasis. Tumori 93: 580-586, 2007.
28. Chok KS and Law WL: Prognostic factors affecting survival and recurrence of patients with pT1 and pT2 colorectal cancer. World J Surg 31: 1485-1490, 2007.
29. Labastida R, Dexeus S, Fábregas R, Tresserra F and Fernández A: Endometrial cancer: Factors affecting survival. Eur J Gynaecol Oncol 24: 381-383, 2003.
30. Kojima F, Yamamoto K, Matsuoka K, Ueda M, Hamada H, Imanishi N and Miyamoto Y: Factors affecting survival after lobectomy with pulmonary artery resection for primary lung cancer. Eur J Cardiothorac Surg 40: e13-e20, 2011.
31. de Magalhães JP: How ageing processes influence cancer. Nat Rev Cancer 13: 357-365, 2013.
32. May M: Statistics: Attacking an epidemic. Nature 509 (Suppl): S50-S51, 2014.
33. Eveno C, Parc Y, Laurent A, Ducreux M and Pocard M: Body-mass index, cancer, and implications for screening. Lancet Oncol 16: e102-e103, 2015.
34. Warren JL, Harlan LC, Trimble EL, Stevens J, Grimes M and Cronin KA: Trends in the receipt of guideline care and survival for women with ovarian cancer: A population-based study. Gynecol Oncol 145: 486-492, 2017.
35. Jensen JS, Jensen DH, Grønhøj C, Karnov KKS, Nørregaard C, Agander TK, Specht L and von Buchwald C: Incidence and survival of oropharyngeal cancer in Denmark: A nation-wide, population-based study from 1980 to 2014. Acta Oncol 57: 269-275, 2018.

36. Welch HG, Schwartz LM and Woloshin S: Are increasing 5-year survival rates evidence of success against cancer? JAMA 283: 2975-2978, 2000.
37. Hallahan C: Data Analysis using SAS. Sociological Methods Res 23: 373-391, 1995.
38. Xu M, Fralick D, Zheng JZ, Wang B, Tu XM and Feng C: The differences and similarities between Two-sample T-Test and Paired T-Test. Shanghai Arch Psychiatry 29: 184-188, 2017.
39. Kim HY: Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test. Restor Dent Endod 42: 152-155, 2017.
40. Cuff J and Higgins JP: Statistical analysis of surgical pathology data using the R program. Adv Anat Pathol 19: 131-139, 2012.
41. Schwartz L, Supuran CT and Alfarouk KO: The Warburg effect and the hallmarks of cancer. Anticancer Agents Med Chem 17: 164-170, 2017.
42. Hofseth LJ and Wargovich MJ: Inflammation, cancer, and targets of ginseng. J Nutr 137 (Suppl 1): 183S-185S, 2007.
43. Cao B, Bray F, Beltran-Sanchez H, Ginsburg O, Soneji S and Soerjomataram I: Benchmarking life expectancy and cancer mortality: Global comparison with cardiovascular disease 1981-2010. BMJ 357: j2765, 2017.
44. May M, Bastian PJ, Brookman-May S, Fritsche HM, Tilki D, Otto W, Bolenz C, Gilfrich C, Trojan L, Herrmann E, et al: Gender-specific differences in cancer-specific survival after radical cystectomy for patients with urothelial carcinoma of the urinary bladder in pathologic tumor stage T4a. Urol Oncol 31: 1141-1147, 2013.
45. Corbett JB and Mori M: Gender-specific cancers, Gender-specific reporters? Twenty-Four years of network TV coverage. Sci Communication 20: 395-408, 1999.
46. Koyama A, Matsuoka H, Ohtake Y, Makimura C, Sakai K, Sakamoto R and Murata M: Gender differences in cancer-related distress in Japan: A retrospective observation study. Biopsychosoc Med 10: 10, 2016.
47. Key T, Appleby P, Barnes I and Reeves G; Endogenous Hormones and Breast Cancer Collaborative Group: Endogenous sex hormones and breast cancer in postmenopausal women: Reanalysis of nine prospective studies. J Natl Cancer Inst 94: 606-616, 2002.
48. Hall SA, Araujo AB, Kupelian V, Maserejian NN and Travison TG: Testosterone and breast cancer. J Sex Med 7: 1035-1037, 2010.
49. Naugler WE, Sakurai T, Kim S, Maeda S, Kim K, Elsharkawy AM and Karin M: Gender disparity in liver cancer due to sex differences in MyD88-dependent IL-6 production. Science 317: 121-124, 2007.
50. NCD Risk Factor Collaboration (NCD-RisC): Trends in adult body-mass index in 200 countries from 1975 to 2014: A pooled analysis of 1698 population-based measurement studies with 19.2 million participants. Lancet 387: 1377-1396, 2016.
51. Bosy-Westphal A, Reinecke U, Schlörke T, Illner K, Kutzner D, Heller M and Müller MJ: Effect of organ and tissue masses on resting energy expenditure in underweight, normal weight and obese adults. Int J Obes Relat Metab Disord 28: 72-79, 2004.
52. Dirat B, Bochet L, Dabek M, Daviaud D, Dauvillier S, Majed B, Wang YY, Meulle A, Salles B, Le Gonidec S, et al: Cancer-associated adipocytes exhibit an activated phenotype and contribute to breast cancer invasion. Cancer Res 71: 2455-2465, 2011.
53. Omabe M, Ezeani M and Omabe KN: Lipid metabolism and cancer progression: The missing target in metastatic cancer treatment. J Appl Biomed 13: 47-59, 2015.
54. Freitas R Jr, Freitas NM, Curado MP, Martins E, Silva CM, Rahal RM and Queiroz GS: Incidence trend for breast cancer among young women in Goiania, Brazil. Sao Paulo Med J 128: 81-84, 2010.
55. Wiener HG, Klinkhamer P, Schenck U, Arbyn M, Bulten J, Bergeron C and Herbert A: European guidelines for quality assurance in cervical cancer screening: Recommendations for cytology laboratories. Cytopathology 18: 67-78, 2007.
56. Arnautovic A, Billups C, Broniscer A, Gajjar A, Boop F and Qaddoumi I: Delayed diagnosis of childhood low-grade glioma: Causes, consequences, and potential solutions. Child Nerv Syst 31: 1067-1077, 2015.
57. Clegg LX, Li FP, Hankey BF, Chu K and Edwards BK: Cancer survival among US whites and minorities: A SEER (Surveillance, Epidemiology, and End Results) Program population-based study. Arch Intern Med 162: 1985-1993, 2002.
58. Weinstock C, Bigenwald R, Hochman T, Sun P, Narod SA and Warner E: Outcomes of surveillance for contralateral breast cancer in patients less than age 60 at the time of initial diagnosis. Curr Oncol 19: e160-e164, 2012.
59. Boyd AL, Reid JC, Salci KR, Aslostovar L, Benoit YD, Shapovalova Z, Nakanishi M, Porras DP, Almakadi M, Campbell CJV, et al: Acute myeloid leukaemia disrupts endogenous myelo-erythropoiesis by compromising the adipocyte bone marrow niche. Nat Cell Biol 19: 1336-1347, 2017.
60. Casartelli G, Dorcaratto A, Ravetti JL, Sola S, Vitali A, Merlo DF and Frosina G: Survival of high grade glioma patients depends on their age at diagnosis. Cancer Biol Ther 8: 1719-1721, 2009.