<u>Original Paper</u>

The Cancer Research Database (CRDB): Integrated Platform to Gain Statistical Insight Into the Correlation Between Cancer and COVID-19

Shahid Ullah[1*], PhD; Farhan Ullah[1], PhD; Wajeeha Rahman[1], PhD; Dimitrios A Karras[2], PhD; Anees Ullah[3], MBBS; Gulzar Ahmad[1], MPhil; Muhammad Ijaz[1], MS; Tianshun Gao[4*], PhD

[1]S-Khan Lab Mardan, Khyber Pakhtunkhwa, Pakistan

[2]Department General, Faculty of Science, National and Kapodistrian University of Athens, Athens, Greece

[3]Kyrgyz State Medical University, Bishkek, Kyrgyzstan

[4]Research Center, The Seventh Affiliated Hospital of Sun Yat-sen University, Shenzhen, China

[*]these authors contributed equally

Corresponding Author:
Tianshun Gao, PhD
Research Center
The Seventh Affiliated Hospital of Sun Yat-sen University
No 628, Zhenyuan Rd
Guangming (New) Dist
Shenzhen, 511464
China
Phone: 86 18126408738
Email: gts.hust@gmail.com

Abstract

Background: The advancement of cancer research has been facilitated through freely available cancer literature, databases, and tools. The age of genomics and big data has given rise to the need for cooperation and data sharing in order to make efficient use of this new information in the COVID-19 pandemic. Although there are many databases for cancer research, their access is not easy owing to different ways of processing and managing the data. There is an absence of a unified platform to manage all of them in a transparent and more comprehensible way.

Objective: In this study, an improved integrated cancer research database and platform is provided to facilitate a deeper statistical insight into the correlation between cancer and the COVID-19 pandemic, unifying the collection of almost all previous published cancer databases and defining a model web database for cancer research, and scoring databases on the basis of the variety types of cancer, sample size, completeness of omics results, and user interface.

Methods: Databases examined and integrated include the Data Portal database, Genomic database, Proteomic database, Expression database, Gene database, and Mutation database; and it is expected that this launch will sort, save, advance the understanding and encourage the use of these resources in the cancer research environment.

Results: To make it easy to search valuable information, 85 cancer databases are provided in the form of a table, and a database of databases named the Cancer Research Database (CRDB) has been built and presented herein. Furthermore, the CRDB has been herein equipped with unique navigation tools in order to be explored by three methods; that is, any single database can be browsed by typing the name in the given search bar, while all categories can be browsed by clicking on the name of the category or image expression icon, thus serving as a facility that could provide all the category databases on a single click.

Conclusions: The computational platform (PHP, HTML, CSS, and MySQL) used to build CRDB for the cancer scientific community can be freely investigated and browsed on the internet and is planned to be updated in a timely manner. In addition, based on the proposed platform, the status and diagnoses statistics of cancer during the COVID-19 pandemic have been thoroughly investigated herein using CRDB, thus providing an easy-to-manage, understandable framework that mines knowledge for future researchers.

XSL•FO
RenderX

KEYWORDS

cancer database; COVID-19; CRDB; genomics; PHP; CRDB; cancer and COVID-19; cancer statistics; cancer research; health database; research platform

Introduction

Cancer is a category of diseases causing irregular cell growth with the ability to infiltrate or spread to other areas of the body. As of 2019, approximately 18 million new cases are reported per year [1], among which 22% of cancer deaths are caused by tobacco use and 10% are caused by obesity, an unhealthy diet, a lack of physical activities, or excessive alcohol use [2]. In the past 2 or 3 decades, recent data have shown that approximately 5%-10% of cancers are caused by genetic disorders [3]. Patients with cancer seem to exhibit exacerbated conditions and a higher mortality rate when exposed to the virus [4]. The COVID-19 pandemic has spread over the world. As in 18 October 2021, there have been 219 million confirmed cases and 4.55 million deaths worldwide, with the number of cases continuing to rise in 216 countries [5]. Patients with cancer are thought to be particularly prone to SARS-CoV-2 infection and the development of more severe COVID-19 symptoms, which could be related to a systemic immunosuppressive condition caused directly by tumor growth and indirectly by anticancer therapy's side effects [6]. Infection can affect people of various ages, but in most situations, disease severity is linked to age limit and pre-existing disorders that decrease immunity, such as cancer. COVID-19 has been linked to an increased risk of severe sickness and death among patients with cancer, according to several studies [7]. Brunello et al [8] suggested that people with cancer are faced with two challenges that potentially lead to death: one is from getting cancer and the other is COVID-19—the latter resulting from undertreatment or overtreatment conditions. Initial investigations revealed that patients with cancer were more likely to contract the virus and become infected with COVID-19.

Because of the effects of antineoplastic therapy, supportive drugs including steroids, and the immunosuppressive qualities of cancer, people with cancer could be immunocompromised [5]. The advancement of modern genomic technology—such as microarrays, proteomics, transcriptomics, and gene sequencing—and the serious situation resulting from the COVID-19 pandemic have resulted in the generation of a huge amount of data [9,10]; therefore, the first challenge of these multi-omics cancer and COVID-19 data was the design and usage of electronic databases to store and manage the large amount of knowledge [11]. A number of databases have been published in this research area, which have explained the wide-ranging information about cancer research and COVID-19 [10,12-16], such as The Cancer Genome Atlas (TCGA), RespCanDB, cBioPortal, Co-19PDB, and ICGC, these databases have been exploring, analyzing, and visualizing multidisciplinary genomics data. Further, we have presented a comparison table with previously published work, in which we have noticed a considerable growth in the number of databases, and we have provided the list of all the cancer databases and have built a database of the databases named the Cancer Research Database (CRDB), with the improvement highlighted in Table 1. To make it easier for researchers and the scientific community, a well-organized and easily accessible platform is required, where all cancer research data can be accessed with a single click. To that end, we have gathered almost all cancer databases and classified them into six categories based on data types: Data Portal database, Genomic database, Proteomic database, Expression database, Gene database, and Mutation database; this would provide an easy way to search data, and users can directly type the name of the needed databases in the search bar or can click the required category, which will lead them to all the databases with a single click. In addition, we have obtained deep insight into the link between cancer and the COVID-19 pandemic, having explained up and down of cancer, new cases, death ratios, etc, before and during the COVID-19 pandemic.

Table 1. Comparison of the Cancer Research Database with other published work.

Database	Databases, n	Type	Year	Component	Journal	Reference
Cancer research database	98	Database+list	2022	Cancer	N/A[a]	N/A
Munich Information Center for Protein Sequences	22	List	2011	Different categories	*Nucleic Acids Research*	[17]
No name	6	Database	2018	Aging	*Nucleic Acids Research*	[18]
COVID-19 pandemic database	59	Database+list	2021	COVID-19	*Computer Methods and Programs in Biomedicine Update*	[10]
Swiss Institute of Bioinformatics	12	Database	2016	Different categories	*Nucleic Acids Research*	[19]
Human cancer databases	58	List	2015	Cancer	*Oncology Reports*	[20]
No name	38	List	2014	Hepatology	*Journal of Hepatology*	[21]
LiverAtlas	53	Databases	2013	Liver	*Liver International*	[22]
No name	16	List	2015	Cancer	*Genomics, Proteomics & Bioinformatics*	[3]

[a]N/A: not applicable.

Previously, we have published several articles in well-known journals, such as the database of Phospho-sites in Animals and Fungi [23] in *Scientific Reports*, the Circadian Gene Database [24] in *Nucleic Acids Research*, Co-19PDB [10] in *Computer Methods and Programs in Biomedicine Update*, DataBases relevant to Human Research [25] in *Future Science* and DataBase of Plant Research [26]; we have provided 15 databases to the scientific community during the COVID-19 pandemic, which can be accessed on the internet [27].

Methods

Construction of the CRDB and Content

We integrated the data from multiple different sources including PubMed, Google, Google Scholar, etc. We used various keywords such as "Cancer database," "cancer database list," and "database of cancer" as search terms to retrieve published cancer-related databases with the help of PubMed. To circumvent missing data, we have manually collected the latest cancer databases from *Nucleic Acids Research*, and *Genomics, Proteomics & Bioinformatics*, which are the leading journals on the database issue. We only collected all cancer databases and have removed all nonfunctional links and programming platforms such as PHP, MySQL, HTML, CSS, and JavaScript have been used to construct the CRDB. Figure 1 shows all the procedures of our database. Finally, we provided a compressive cancer research database to the scientific community, which is easy to operate and will be updated over time.

Figure 1. Flowchart and procedure for the collection and integration of cancer databases and the construction of the Cancer Research Database (CRDB).
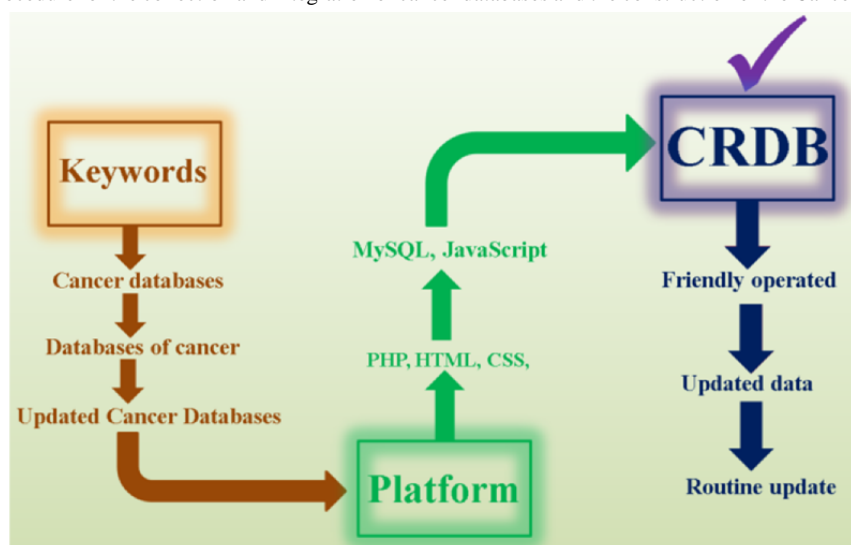
Database Classification

Several articles have been published in this research area [28-30], each has its their own classification of databases based on its function, application, technological feature, and organism, such as human and mouse [29], Plant [26], Drosophila [31], fungi, COVID-19, etc. According to such published works, we

have also classified the cancer databases into six categories: Data Portal database, Genomic database, Proteomic database, Expression database, Gene database, and Mutation database—their details are given bellow.

Expression Databases

In cancer expression databases, the expression levels of thousands of genes can be continuously measured under particular experimental environments and conditions resulting from marked advancements in DNA microarray technology. This technology made it possible to understand life at the molecular level, and enables us to generate large-scale gene expression data. It has also been applied in a wide range of applications such as cancer prediction, diagnosis, and drug discovery, which are very important issues for cancer treatment [32]. Some well-known expression databases including BioXpress [33], miRCancer [34], and Gene Expression Database [35] are shown in Figure 2A.

Figure 2. Main pages of some commonly using cancer databases. (A) A screenshot of the Expression Database named GXD, (B) a screenshot of the Data Portal category named CNVs (copy number variations), (C) "IARCTP53": a database of the Gene category, (D, E, and F) main pages of Proteomic database, Mutation database, and Genomic database respectively.

Data Portal Databases

Data Portal is a type of database that provides comprehensive genomic, epigenomic, transcriptomic, and proteomic data. a large number of data are publicly available for anyone in the research community and are used to diagnose, treat, and prevent cancer [12]. There are different published databases such as The European Genome-phenome Archive (EGA), which is a data center for all types of sequencing and genotyping experiments. Almost 58% of all studies in the EGA are related to cancer [36]. The "CanEvolve" database fulfills the need for data integration and interpretation. It contains data from 90 studies involving more than 10,000 patients. Data analysis can be performed at different levels: primary analysis including mRNA, microRNA (miRNA), and protein expression, genome variations, and protein–protein interactions; integrative analysis of gene and miRNA expression, gene expression, and copy number variations, and gene set enrichment analysis; network analysis; and survival analysis [37]; the main page of this database is shown in Figure 2B.

Gene Databases

Gene databases collect various types of gene data and information related to cancer [38] and also provide help to researchers in understanding the genetic architecture of complex diseases and improve the accuracy of diagnosis and the effectiveness of therapy [39]. Various databases have been published such as the "IARCTP53" shown in Figure 2C; this database compiles various types of data and information on human *TP53* variations related to cancer [38]. The "TGDBs" gene database includes mechanisms of oncogenic activation, regulation, frequency of involvement in various tumor types, and chromosomal location. Data about the encoded proteins includes the cell type in which they are found, subcellular location, DNA-, protein-, and ligand-binding, role in development, and normal biochemical function [40].

Proteomic Databases

Cancer proteome databases encompass tumor tissues, cells, and biological fluids to interpret signaling pathways, identify signatures related to tumor initiation, invasion, and metastasis, and determine analytical, predictive, and prognostic markers [41], and also help determine the molecular details of proteome differentiation in various human tissues and organs, thus greatly improving our understanding of disease and human biology [23,42]. A number of databases have been published in this research area such as those shown in Figure 2D: "CanProVar" is designed to store and display single amino acid alterations including both germline and somatic variations in the human proteome, particularly those related to the genesis or development of human cancer based on the published studies and sources [43].

Mutation Databases

Mutation databases play an important role in science, diagnostics, and genetic health care and can play a vital role in life and death decisions. These databases are extensively used, but only gene- or locus-specific databases have been previously reviewed for their utility, accuracy, completeness, and currency [44]. Mutations in the tumor suppressor gene *TP53* are associated with a variety of cancers [45]. More than 50% of the human tumors harbor *TP53* mutations, resulting in a collection of over 45,000 somatic and germline mutations in the UMDTP53 database, as shown in Figure 2E. Analyses of these mutations have been helpful for improving our knowledge on the structure-function relations within the TP53 protein [46]. A number of databases have been published in this research area, which are of marked utility in the scientific community.

Genomic Databases

The Cancer Genome Database represents one of numerous international groups dedicated to performing wide-ranging genomic and epigenomic studies of selected cancer types to develop our understanding of disease and provide an open-access resource for international cancer research [47]. This database is aimed at improving the understanding of the molecular basis of cancer development [48]. Several databases have been published in this research area; for example, the MethCNA comprehensive database for genomic data in human cancer (Figure 2F). Per a most recent publication, this database contains approximately 10,000 tumor samples covering 37 cancer types. All the data were collected from the TCGA and the National Center for Biotechnology Information and were evaluated using a pipeline that combined multiple computational resources and tools [49]. BioMuta is a single-nucleotide variation and disease association database where variations are mapped to genomes and RefSeq nucleotide entries and are incorporated through UniProtKB/Swiss-Prot positional coordinates. The recent version of BioMuta contains only nonsynonymous single-nucleotide variations associated with cancer [50].

Results and Discussion

Database Statistics

In this work, we have provided almost all cancer databases (Table S1 in Multimedia Appendix 1) and shown the year-wise growth of the CRDB. Figure 3 shows the magnitude of category-wise growth of the databases. Table 2 shows the year-wise growth distribution of cancer databases, which marks tremendous growth and is an achievement for the cancer scientific community. Further, we have modified or deleted all the nonfunctional and inaccessible database links and have provided a new, updated cancer database in the form of a database named CRDB and a table (Table S1 in Multimedia Appendix 1).

Figure 3. The statistics data of the Cancer Research Database (DB)—distribution of the database category.
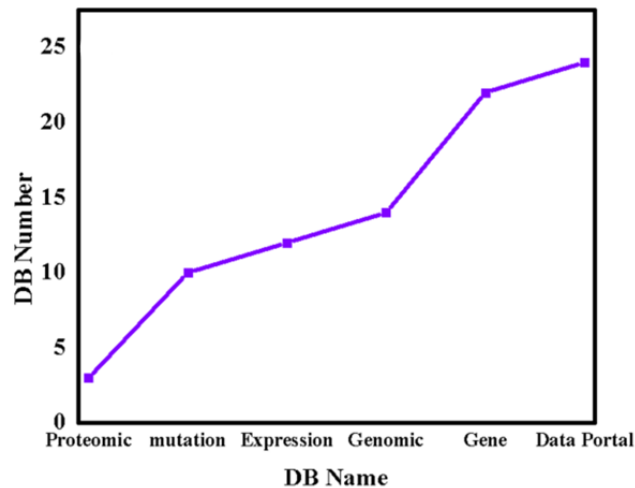
**Table 2.** Year-wise growth of the Cancer Research Database.

Year	Database growth, %
1999	1
2004	1
2005	1
2006	2
2007	2
2008	2
2009	2
2010	5
2011	6
2012	2
2013	7
2014	5
2015	6
2016	6
2017	11
2018	7
2019	7
2020	14
2021	12

Usage of the CRDB

The CRDB has been developed to provide an easy and user-friendly search experience; for easier and faster search, three options are provided for finding cancer databases. First, browsing can be carried out by typing the name of the database in the search bar, which is highlighted in Figure 4A, or by clicking on the name of the category or image expression, which is shown in Figure 4B. with CRDB statistics as well, which will lead to the category list page (Figure 4C), and a brief overview with the original link of the required search will be accessed by clicking the needed database. Further, for a specific database search, the "CT Database" is used as an example from the expression Databases to make it more user-friendly.

Figure 4. Browse options of the Cancer Research Database (CRDB). (A) Can be browsed by typing the name. (B) Can be browsed by category name or image expression. (C) An example and the final result.
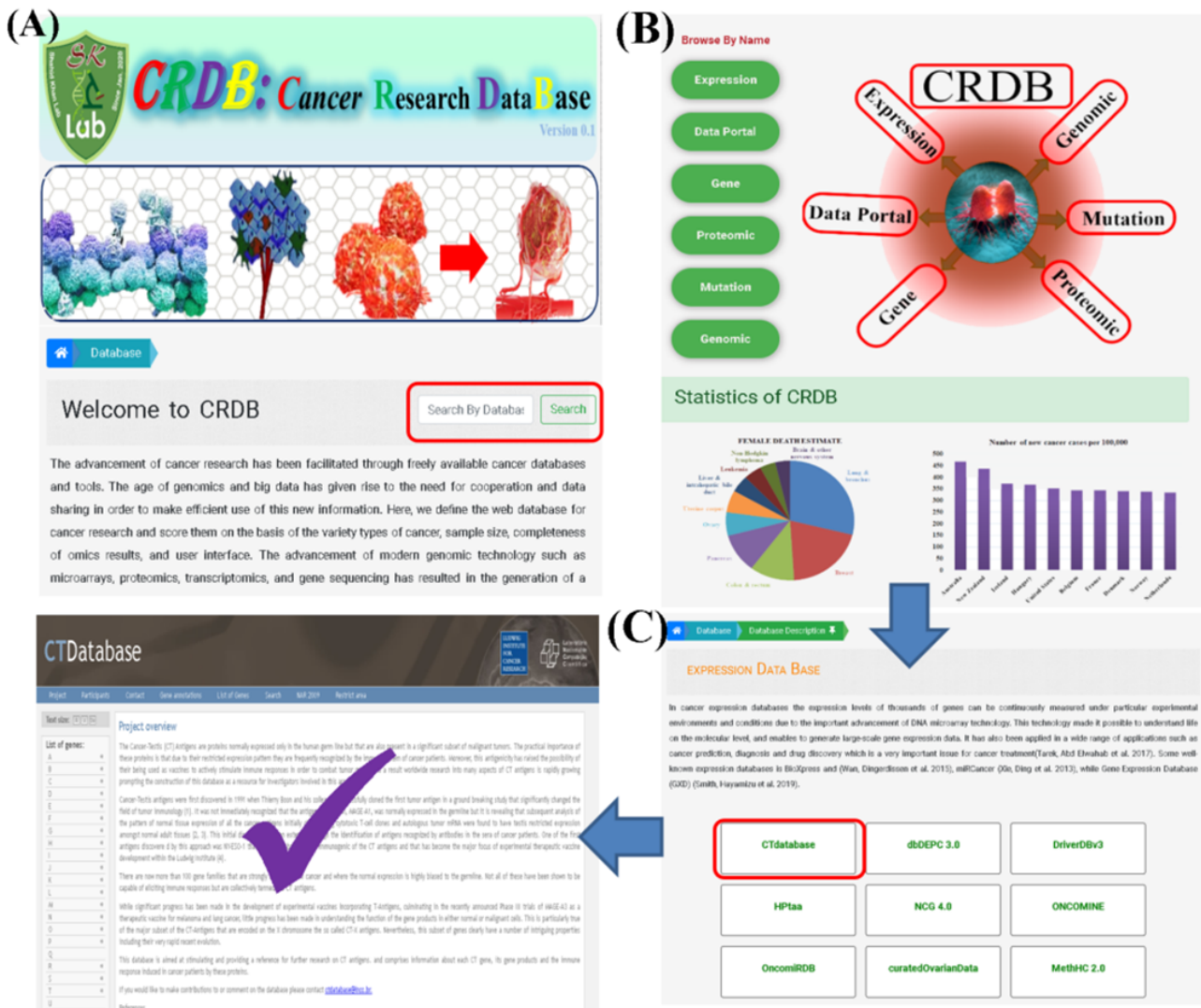
Example of Cancer Diagnosis During the COVID-19 Pandemic

As of the previously reported reductions in cancer screening and other preventive care visits during the COVID-19 pandemic, the number of new cancer cases in 2020 is likely to be smaller than anticipated. According to one survey of diagnostic results, there was a 46% decrease in diagnosis of six different cancers (colorectal, pancreatic breast, lung, esophageal, and stomach cancer) from March 1 to April 18, 2020, relative to the period between January 6, 2019, and February 29, 2020, varying from a 25% decrease in the detection rate of pancreatic cancer to a 52% decrease in that of breast cancer [51,52]. Another study found that new CRC diagnoses were 30% lower from January to mid-April 2020 relative to the same timeframe in 2019 [53]. Across the world, similar losses have been noted, including those in the United Kingdom [54], the United States [54], and the Netherlands [55]. While these preliminary observations may provide insight into the pandemic's effect on cancer diagnosis, population-based cancer registry evidence and the degree to which these delays may lead to more advanced-stage disease will not be available for some time.

COVID-19 and Cancer

People with active cancer are more vulnerable to infectious pathogens as a result of a compromised immune system due to the malignancy and its treatment (eg, surgery and chemotherapy). This has raised fears that COVID-19–related problems and mortality may be more common among patients with cancer [56]. According to a 2020 study, patients with cancer may be at a higher risk of COVID-19 than those without cancer [57]. COVID-19 infection can impact persons with a wide range of hematologic diseases; however, the risk of infection is lower in patients with chronic myeloid proliferative neoplasms such as chronic myeloid leukemia and greater in persons on immunosuppressive medication [57,58]. A study at a tertiary care hospital in Wuhan, China, reported that patients with lung cancer above the age 60 years are at a high risk of COVID-19 locally [14,58,59] and worldwide [59,60]. It was also revealed that of the many cancer types, people with lung cancer who are over 60 years old are especially susceptible to COVID-19. Although it may seem intuitive that people with a defective respiratory epithelium are more susceptible to rapid virus entry into the lungs [61]. Another study shows a decrease in 6 types of cancer, with the total number of detected cancers

being 4310 before and 2310 during the COVID-19 pandemic, with breast cancer ranking the highest with 2208 cases before and 2310 cases during the COVID-19 pandemic, followed by colorectal cancer at 946 cases before and 840 cases during the pandemic, and Figure 5 shows details regarding all 6 cancer databases before and during the COVID-19 pandemic [52].

**Figure 5.** The number of cancers detected before and during the COVID-19 pandemic.
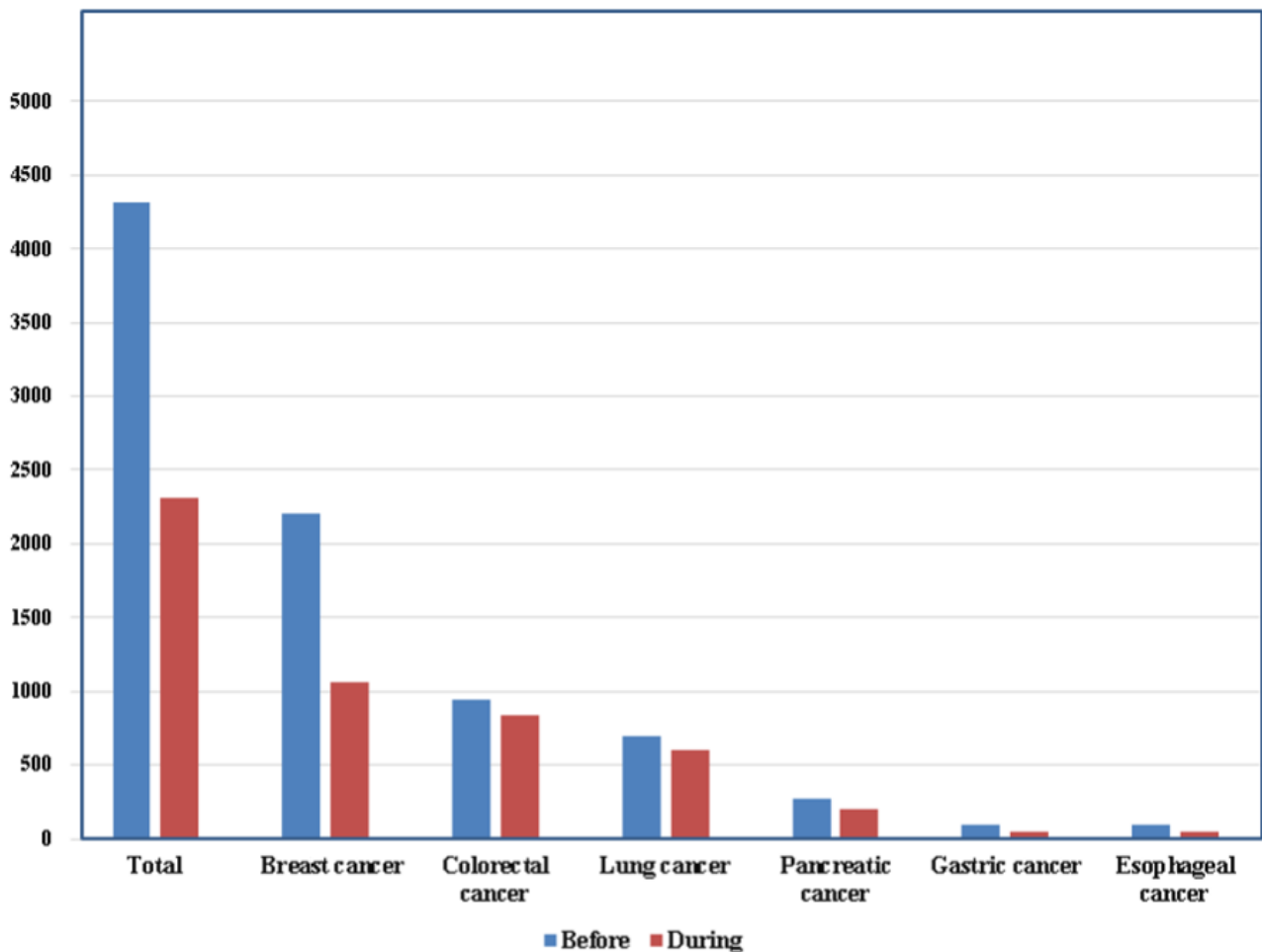
Global Cancer Rate

According to American institute for cancer research [62], the rates of all cancers have increased in almost in all countries, the privation of which is the major task for public health in 2021. Decreasing the cancer rate involves coordinated and comprehensive intervention from all facets of society, particularly the public sphere, civil society, and health and other occupations. Figure 6 shows the top 10 country-wise cancer rate from Oceania, Europe, and North America, in which the highest rate of cancer was reported in Australia (468.0 people per 100,000 population). For these 10 countries, the age-standardized average was at least 320 people per 100,000 population. With 579.9 men per 100,000 population, the age-standardized average was at least 360 people per 100,000 population, and with 363.0 women per 100,000 population, the age-standardized average was at least 300 people per 100,000 population.

Figure 6. Country-wise cancer rate with an age-standardized average (per 100,000 population).
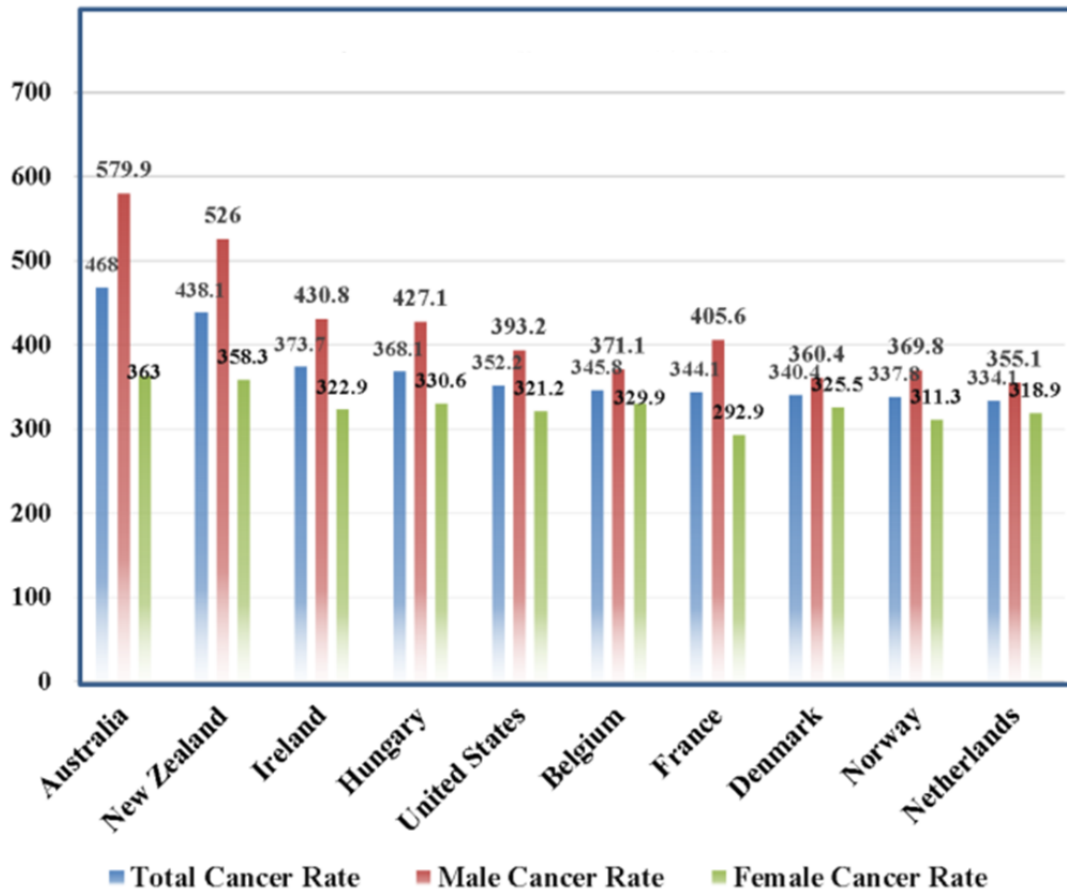
New Cases and Deaths

The American Cancer Society Cancer Action NetworkSM works worldwide to increase the quality of care for patients with and survivors of cancer. With time and the emergence of new cases worldwide, we have compiled a list of the top 10 cancers diagnosed in the United States in 2021. Table 3 shows the number of new cancer cases, with breast cancer in women and prostate cancer in men ranking first (30% cases) and second (26%), respectively. Although mortality estimates are shown in Table 3, lung and bronchial cancer showed the same rates in both male and female patients and ranked the highest, followed by breast and prostate cancer.

Table 3. Rates of new cancers and mortality between male and female patients.

Cancer type	New cancers, %		Mortality, %	
	Male	Female	Male	Female
Prostate	26	N/A[a]	11	N/A
Breast	N/A	30	N/A	15
Lung and bronchial	12	13	22	22
Colorectal	8	8	9	8
Urinary bladder	7	—[b]	4	—
Skin melanoma	6	5	—	—
Kidney and renal pelvis	5	3	—	—
Non-Hodgkin lymphoma	5	4	4	3
Oral cavity and pharynx	4	—	—	—
Leukemia	4	3	4	3
Pancreatic	3	3	8	8
Uterine corpus	N/A	7	N/A	4
Brain and other nervous system regions	—	—	3	3
Liver and intrahepatic bile duct	—	—	6	3
Esophageal	—	—	4	—
Ovarian	N/A	—	N/A	5

[a]N/A: not applicable.

[b]—: not determined.

Conclusions

A biological database provides facilities for storing, organizing, and retrieving biological data such as DNA, RNA, carbohydrates, proteins, and cancers. It can be easily viewed, managed, and modified. A number of papers have been published in this research field, which have their own classification of cancer databases based on their function, use, certain technical aspects, and on species such as human, mouse, plant, and fungi. According to such published studies, we have classified the cancer databases into six categories: Data Portal database, Genomic database, Proteomic database, Expression database, Gene database, and Mutation database. Further, we have collected almost all cancer databases with a short introduction and have updated or removed all nonfunctional links. Furthermore, we have understood the current situation of cancer and its correlation with COVID-19; for example, the up-down, mortality, and new case count based on continent and countries, etc. Our database can be searched through an easy-to-use, user-friendly method, can be searched by clicking on category name of image expression, or users can type the name of needed databases in the given search bar and they will be updated with time. In addition, we have examined the status and diagnoses of cancer during the COVID-19 pandemic and have provided easy and understandable information for future researchers.

Availability of Data and Material

These data will be available under the journal rule and regulation. To avoid future conflict and plagiarism issue, CRDB database is uploaded on the internet [63] so that we have provided some content in this article.

Authors' Contributions

SU supervised the study with TG, DAK, WR, GA, FU, MI, AU and collected and verified the data carefully. SU drafted the manuscript. All authors reviewed the manuscript and agreed to submit it. Both TG and SU are the corresponding authors for this manuscript.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1

Cancer databases.

[DOCX File , 22 KB-Multimedia Appendix 1]

References

1. Sciacovelli M, Schmidt C, Maher ER, Frezza C. Metabolic Drivers in Hereditary Cancer Syndromes. Annu Rev Cancer Biol 2020 Mar 09;4(1):77-97. [doi: 10.1146/annurev-cancerbio-030419-033612]

2. Jayasekara H, MacInnis RJ, Room R, English DR. Long-Term Alcohol Consumption and Breast, Upper Aero-Digestive Tract and Colorectal Cancer Risk: A Systematic Review and Meta-Analysis. Alcohol Alcohol 2016 May;51(3):315-330. [doi: 10.1093/alcalc/agv110] [Medline: 26400678]

3. Yang Y, Dong X, Xie B, Ding N, Chen J, Li Y, et al. Databases and web tools for cancer genomics study. Genomics Proteomics Bioinformatics 2015 Feb;13(1):46-50 [FREE Full text] [doi: 10.1016/j.gpb.2015.01.005] [Medline: 25707591]

4. Han HJ, Nwagwu C, Anyim O, Ekweremadu C, Kim S. COVID-19 and cancer: From basic mechanisms to vaccine development using nanotechnology. Int Immunopharmacol 2021 Jan;90:107247 [FREE Full text] [doi: 10.1016/j.intimp.2020.107247] [Medline: 33307513]

5. Tian J, Miao X. Challenges and recommendations for cancer care in the COVID-19 pandemic. Cancer Biol Med 2020 Aug 15;17(3):515-518 [FREE Full text] [doi: 10.20892/j.issn.2095-3941.2020.0300] [Medline: 32944386]

6. Liu C, Zhao Y, Okwan-Duodu D, Basho R, Cui X. COVID-19 in cancer patients: risk, clinical features, and management. Cancer Biol Med 2020 Aug 15;17(3):519-527 [FREE Full text] [doi: 10.20892/j.issn.2095-3941.2020.0289] [Medline: 32944387]

7. Pathania AS, Prathipati P, Abdul BA, Chava S, Katta SS, Gupta SC, et al. COVID-19 and Cancer Comorbidity: Therapeutic Opportunities and Challenges. Theranostics 2021;11(2):731-753 [FREE Full text] [doi: 10.7150/thno.51471] [Medline: 33391502]

8. Brunello A, Galiano A, Finotto S, Monfardini S, Colloca G, Balducci L, et al. Older cancer patients and COVID-19 outbreak: Practical considerations and recommendations. Cancer Med 2020 Dec;9(24):9193-9204 [FREE Full text] [doi: 10.1002/cam4.3517] [Medline: 33219746]

9. Sarode GS, Sarode SC, Maniyar N, Anand R, Patil S. Oral cancer databases: A comprehensive review. J Oral Pathol Med 2018 Jul;47(6):547-556. [doi: 10.1111/jop.12667] [Medline: 29193424]

10. Ullah S, Ullah A, Rahman W, Ullah F, Khan SB, Ahmad G, et al. An innovative user-friendly platform for Covid-19 pandemic databases and resources. Comput Methods Programs Biomed Update 2021;1:100031 [FREE Full text] [doi: 10.1016/j.cmpbup.2021.100031] [Medline: 34604832]

11. Percudani R, Carnevali D, Puggioni V. Ureidoglycolate hydrolase, amidohydrolase, lyase: how errors in biological databases are incorporated in scientific papers and vice versa. Database (Oxford) 2013;2013:bat071 [FREE Full text] [doi: 10.1093/database/bat071] [Medline: 24107613]

12. Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, Mills GB, Shaw KRM, Ozenberger BA, et al. The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet 2013 Oct;45(10):1113-1120 [FREE Full text] [doi: 10.1038/ng.2764] [Medline: 24071849]

13. Zhang J, Baran J, Cros A, Guberman JM, Haider S, Hsu J, et al. International Cancer Genome Consortium Data Portal--a one-stop shop for cancer genomics data. Database (Oxford) 2011;2011:bar026 [FREE Full text] [doi: 10.1093/database/bar026] [Medline: 21930502]

14. Ullah S. EDBCO-19: Emergency Data Base of COVID-19. J Clin Med Res 2020 Aug 1. [doi: 10.37191/mapsci-2582-4333-2(5)-047]

15. Bakouny Z, Hawley J, Choueiri T, Peters S, Rini B, Warner J, et al. COVID-19 and Cancer: Current Challenges and Perspectives. Cancer Cell 2020 Nov 09;38(5):629-646 [FREE Full text] [doi: 10.1016/j.ccell.2020.09.018] [Medline: 33049215]

16. International Cancer Genome Consortium T, Hudson TJ, Anderson W, Artez A, Barker AD, Bell C, et al. International network of cancer genome projects. Nature 2010 Apr 15;464(7291):993-998 [FREE Full text] [doi: 10.1038/nature08987] [Medline: 20393554]

17. Mewes HW, Ruepp A, Theis F, Rattei T, Walter M, Frishman D, et al. MIPS: curated databases and comprehensive secondary data resources in 2010. Nucleic Acids Res 2011 Jan;39(Database issue):D220-D224 [FREE Full text] [doi: 10.1093/nar/gkq1157] [Medline: 21109531]

18. Tacutu R, Thornton D, Johnson E, Budovsky A, Barardo D, Craig T, et al. Human Ageing Genomic Resources: new and updated databases. Nucleic Acids Res 2018 Jan 04;46(D1):D1083-D1090 [FREE Full text] [doi: 10.1093/nar/gkx1042] [Medline: 29121237]

19. SIB Swiss Institute of Bioinformatics Members. The SIB Swiss Institute of Bioinformatics' resources: focus on curated databases. Nucleic Acids Res 2016 Jan 04;44(D1):D27-D37 [FREE Full text] [doi: 10.1093/nar/gkv1310] [Medline: 26615188]

20. Pavlopoulou A, Spandidos D, Michalopoulos I. Human cancer databases (review). Oncol Rep 2015 Jan;33(1):3-18 [FREE Full text] [doi: 10.3892/or.2014.3579] [Medline: 25369839]

21. Teufel A. Bioinformatics and database resources in hepatology. J Hepatol 2015 Mar;62(3):712-719 [FREE Full text] [doi: 10.1016/j.jhep.2014.10.036] [Medline: 25450718]

22. Zhang Y, Yang C, Wang S, Chen T, Li M, Wang X, et al. LiverAtlas: a unique integrated knowledge database for systems-level research of liver and hepatic disease. Liver Int 2013 Sep;33(8):1239-1248. [doi: 10.1111/liv.12173] [Medline: 23601370]

23. Ullah S, Lin S, Xu Y, Deng W, Ma L, Zhang Y, et al. dbPAF: an integrative database of protein phosphorylation in animals and fungi. Sci Rep 2016 Mar 24;6:23534 [FREE Full text] [doi: 10.1038/srep23534] [Medline: 27010073]

24. Li S, Shui K, Zhang Y, Lv Y, Deng W, Ullah S, et al. CGDB: a database of circadian genes in eukaryotes. Nucleic Acids Res 2017 Jan 04;45(D1):D397-D403 [FREE Full text] [doi: 10.1093/nar/gkw1028] [Medline: 27789706]

25. Ullah S, Rahman W, Ullah F, Ahmad G, Ijaz M, Gao T. DBHR: a collection of databases relevant to human research. Future Sci OA 2022 Mar;8(3):FSO780 [FREE Full text] [doi: 10.2144/fsoa-2021-0101] [Medline: 35251694]

26. Ullah S, Rahman W, Ullah F, Ahmad G, Ijaz M, Gao T. DBPR: DataBase of Plant Research. Research Square. URL: https://www.researchsquare.com/article/rs-153167/v1 [accessed 2022-04-22]

27. Khan SA. Tremendous Contribution of Dr. Shahid Ullah to Scientific Community during COVID-19 Pandemic in the Form of Scientific Research. J Clin Med Res 2020 Aug 1. [doi: 10.37191/mapsci-2582-4333-2(5)-049]

28. Xu D. Protein databases on the internet. Curr Protoc Protein Sci 2012 Nov;Chapter 2:Unit2.6. [doi: 10.1002/0471140864.ps0206s70] [Medline: 23151744]

29. Celis J, Ostergaard M, Jensen N, Gromova I, Rasmussen H, Gromov P. Human and mouse proteomic databases: novel resources in the protein universe. FEBS Lett 1998 Jun 23;430(1-2):64-72 [FREE Full text] [doi: 10.1016/s0014-5793(98)00527-4] [Medline: 9678596]

30. Harper R. Access to DNA and protein databases on the Internet. Curr Opin Biotechnol 1994 Feb;5(1):4-18. [doi: 10.1016/s0958-1669(05)80063-1] [Medline: 7764641]

31. Sanchez C, Lachaize C, Janody F, Bellon B, Röder L, Euzenat J, et al. Grasping at molecular interactions and genetic networks in Drosophila melanogaster using FlyNets, an Internet database. Nucleic Acids Res 1999 Jan 01;27(1):89-94 [FREE Full text] [doi: 10.1093/nar/27.1.89] [Medline: 9847149]

32. Tarek S, Abd Elwahab R, Shoman M. Gene expression based cancer classification. Egypt Inform J 2017 Nov;18(3):151-159. [doi: 10.1016/j.eij.2016.12.001]

33. Wan Q, Dingerdissen H, Fan Y, Gulzar N, Pan Y, Wu T, et al. BioXpress: an integrated RNA-seq-derived gene expression database for pan-cancer analysis. Database (Oxford) 2015;2015 [FREE Full text] [doi: 10.1093/database/bav019] [Medline: 25819073]

34. Xie B, Ding Q, Han H, Wu D. miRCancer: a microRNA-cancer association database constructed by text mining on literature. Bioinformatics 2013 Mar 01;29(5):638-644. [doi: 10.1093/bioinformatics/btt014] [Medline: 23325619]

35. Smith C, Hayamizu T, Finger J, Bello S, McCright IJ, Xu J, et al. The mouse Gene Expression Database (GXD): 2019 update. Nucleic Acids Res 2019 Jan 08;47(D1):D774-D779 [FREE Full text] [doi: 10.1093/nar/gky922] [Medline: 30335138]

36. Lappalainen I, Almeida-King J, Kumanduri V, Senf A, Spalding JD, Ur-Rehman S, et al. The European Genome-phenome Archive of human data consented for biomedical research. Nat Genet 2015 Jul;47(7):692-695 [FREE Full text] [doi: 10.1038/ng.3312] [Medline: 26111507]

37. Samur MK, Yan Z, Wang X, Cao Q, Munshi NC, Li C, et al. canEvolve: a web portal for integrative oncogenomics. PLoS One 2013;8(2):e56228 [FREE Full text] [doi: 10.1371/journal.pone.0056228] [Medline: 23418540]

38. Olivier M, Eeles R, Hollstein M, Khan MA, Harris CC, Hainaut P. The IARC TP53 database: new online mutation analysis and recommendations to users. Hum Mutat 2002 Jun;19(6):607-614. [doi: 10.1002/humu.10081] [Medline: 12007217]

39. Su L, Liu G, Bai T, Meng X, Ma Q. MGOGP: a gene module-based heuristic algorithm for cancer-related gene prioritization. BMC Bioinformatics 2018 Jun 05;19(1):215 [FREE Full text] [doi: 10.1186/s12859-018-2216-0] [Medline: 29871590]

40. Gong X, Wu R, Zhang Y, Zhao W, Cheng L, Gu Y, et al. Extracting consistent knowledge from highly inconsistent cancer gene data sources. BMC Bioinformatics 2010 Feb 05;11:76 [FREE Full text] [doi: 10.1186/1471-2105-11-76] [Medline: 20137077]

41. Arntzen M, Boddie P, Frick R, Koehler CJ, Thiede B. Consolidation of proteomics data in the Cancer Proteomics database. Proteomics 2015 Nov;15(22):3765-3771. [doi: 10.1002/pmic.201500144] [Medline: 26316313]

42. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. Science 2015 Jan 23;347(6220):1260419. [doi: 10.1126/science.1260419] [Medline: 25613900]

43. Li J, Duncan DT, Zhang B. CanProVar: a human cancer proteome variation database. Hum Mutat 2010 Mar;31(3):219-228 [FREE Full text] [doi: 10.1002/humu.21176] [Medline: 20052754]

44. George RA, Smith TD, Callaghan S, Hardman L, Pierides C, Horaitis O, et al. General mutation databases: analysis and review. J Med Genet 2008 Feb;45(2):65-70 [FREE Full text] [doi: 10.1136/jmg.2007.052639] [Medline: 17893115]

45. Li VD, Li KH, Li JT. TP53 mutations as potential prognostic markers for specific cancers: analysis of data from The Cancer Genome Atlas and the International Agency for Research on Cancer TP53 Database. J Cancer Res Clin Oncol 2019 Mar;145(3):625-636. [doi: 10.1007/s00432-018-2817-z] [Medline: 30542790]

46. Leroy B, Anderson M, Soussi T. TP53 mutations in human cancer: database reassessment and prospects for the next decade. Hum Mutat 2014 Jun;35(6):672-688. [doi: 10.1002/humu.22552] [Medline: 24665023]

47. Cooper LA, Demicco EG, Saltz JH, Powell RT, Rao A, Lazar AJ. PanCancer insights from The Cancer Genome Atlas: the pathologist's perspective. J Pathol 2018 Apr;244(5):512-524 [FREE Full text] [doi: 10.1002/path.5028] [Medline: 29288495]

48. Lee J. Exploring cancer genomic data from the cancer genome atlas project. BMB Rep 2016 Nov;49(11):607-611 [FREE Full text] [doi: 10.5483/bmbrep.2016.49.11.145] [Medline: 27530686]

49. Deng G, Yang J, Zhang Q, Xiao Z, Cai H. MethCNA: a database for integrating genomic and epigenomic data in human cancer. BMC Genomics 2018 Feb 13;19(1):138 [FREE Full text] [doi: 10.1186/s12864-018-4525-0] [Medline: 29433427]

50. Dingerdissen H, Torcivia-Rodriguez J, Hu Y, Chang T, Mazumder R, Kahsay R. BioMuta and BioXpress: mutation and expression knowledgebases for cancer biomarker discovery. Nucleic Acids Res 2018 Jan 04;46(D1):D1128-D1136 [FREE Full text] [doi: 10.1093/nar/gkx907] [Medline: 30053270]

51. Biondi‐Zoccai G, Sciarretta S, Bullen C, Nocella C, Violi F, Loffredo L, et al. Acute Effects of Heat‐Not‐Burn, Electronic Vaping, and Traditional Tobacco Combustion Cigarettes: The Sapienza University of Rome‐Vascular Assessment of Proatherosclerotic Effects of Smoking (SUR‐VAPES) 2 Randomized Trial. JAHA 2019 Mar 19;8(6). [doi: 10.1161/jaha.118.010455]

52. Kaufman HW, Chen Z, Niles J, Fesko Y. Changes in the Number of US Patients With Newly Identified Cancer Before and During the Coronavirus Disease 2019 (COVID-19) Pandemic. JAMA Netw Open 2020 Aug 03;3(8):e2017267 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.17267] [Medline: 32749465]

53. Antoniewicz L, Brynedal A, Hedman L, Lundbäck M, Bosson JA. Acute Effects of Electronic Cigarette Inhalation on the Vasculature and the Conducting Airways. Cardiovasc Toxicol 2019 Oct;19(5):441-450 [FREE Full text] [doi: 10.1007/s12012-019-09516-x] [Medline: 30963443]

54. Waziry R, Jawad M, Ballout RA, Al Akel M, Akl EA. The effects of waterpipe tobacco smoking on health outcomes: an updated systematic review and meta-analysis. Int J Epidemiol 2017 Feb 01;46(1):32-43. [doi: 10.1093/ije/dyw021] [Medline: 27075769]

55. Leventhal AM, Strong DR, Kirkpatrick MG, Unger JB, Sussman S, Riggs NR, et al. Association of Electronic Cigarette Use With Initiation of Combustible Tobacco Product Smoking in Early Adolescence. JAMA 2015 Aug 18;314(7):700-707 [FREE Full text] [doi: 10.1001/jama.2015.8950] [Medline: 26284721]

56. Siegel R, Miller K, Fuchs H, Jemal A. Cancer Statistics, 2021. CA Cancer J Clin 2021 Jan;71(1):7-33 [FREE Full text] [doi: 10.3322/caac.21654] [Medline: 33433946]

57. Cattaneo C, Daffini R, Pagani C, Salvetti M, Mancini V, Borlenghi E, et al. Clinical characteristics and risk factors for mortality in hematologic patients affected by COVID-19. Cancer 2020 Dec 01;126(23):5069-5076 [FREE Full text] [doi: 10.1002/cncr.33160] [Medline: 32910456]

58. Yu J, Ouyang W, Chua MLK, Xie C. SARS-CoV-2 Transmission in Patients With Cancer at a Tertiary Care Hospital in Wuhan, China. JAMA Oncol 2020 Jul 01;6(7):1108-1110 [FREE Full text] [doi: 10.1001/jamaoncol.2020.0980] [Medline: 32211820]

59. UK Coronavirus Cancer Monitoring Project team. The UK Coronavirus Cancer Monitoring Project: protecting patients with cancer in the era of COVID-19. Lancet Oncol 2020 May;21(5):622-624 [FREE Full text] [doi: 10.1016/S1470-2045(20)30230-8] [Medline: 32304634]

60. Li H, Liu SM, Yu XH, Tang SL, Tang CK. Coronavirus disease 2019 (COVID-19): current status and future perspectives. Int J Antimicrob Agents 2020 May;55(5):105951 [FREE Full text] [doi: 10.1016/j.ijantimicag.2020.105951] [Medline: 32234466]

61. Zhang H, Wang L, Chen Y, Wu Q, Chen G, Shen X, et al. Outcomes of novel coronavirus disease 2019 (COVID-19) infection in 107 patients with cancer from Wuhan, China. Cancer 2020 Sep 01;126(17):4023-4031 [FREE Full text] [doi: 10.1002/cncr.33042] [Medline: 32573776]

62. Global cancer data by country. World Cancer Research Fund International. URL: https://www.wcrf.org/cancer-trends/global-cancer-data-by-country/ [accessed 2022-04-25]

63. Welcome to CRDB. CRDB: Cancer Research Database. URL: https://habdsk.org/crdb.php [accessed 2022-04-25]

Abbreviations

CNV: copy number variation
CRDB: cancer research database
EGA: European Genome-phenome Archive
miRNA: microRNA
TCGA: The Cancer Genome Atlas

XSL•FO
RenderX