Epidemiology of gallbladder cancer in the United States: a population-based study

Motasem Alkhayyat¹, Mohannad Abou Saleh², Thabet Qapaja¹, Mohammad Abureesh¹, Ashraf Almomani¹, Emad Mansoor⁴, Prabhleen Chahal²

¹Department of Internal Medicine, Cleveland Clinic Foundation, Cleveland, OH, USA; ²Department of Gastroenterology, Hepatology, and Nutrition, Cleveland Clinic Foundation, Cleveland, OH, USA; ³Department of Internal Medicine, Staten Island University Hospital, Staten Island, NY, USA; ⁴Department of Gastroenterology and Hepatology, University Hospitals Cleveland Medical Center, Cleveland, OH, USA

Contributions: (I) Conception and design: M Alkhayyat, M Abou Saleh, E Mansoor, P Chahal; (II) Administrative support: E Mansoor, M Abou Saleh; (III) Provision of study materials or patients: M Alkhayyat, M Abureesh, A Almomani, P Chahal; (IV) Collection and assembly of data: M Alkhayyat, M Abureesh, A Almomani, T Qapaja; (V) Data analysis and interpretation: M Alkhayyat, M Abou Saleh, M Abureesh, A Almomani; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Mohannad Abou Saleh, MD. Gastroenterology, Hepatology and Nutrition, Digestive Disease and Surgery Institute, Cleveland Clinic Foundation, 9500 Euclid Avenue/A30, Cleveland, OH 44195, USA. Email: Abousam@ccf.org.

Background: Gallbladder cancer (GBC) is the most common neoplasm of the biliary tract with the lowest rates of survival. Most GBCs are adenocarcinomas that arise from the epithelial lining of the gallbladder. There are limited data in the literature regarding the epidemiology of GBC. Using a large database, we aim to describe the epidemiology using a US population database.

Methods: A multi-institutional database (Explorys Inc., Cleveland, OH, USA) was surveyed. A cohort of patients with a primary malignant neoplasm of gallbladder between 1999–2019 was identified. The prevalence rate was calculated and age-, race-, and sex-based distributions were described. Multivariate analysis was done to evaluate underlying associations.

Results: Of the 56,197,690 individuals in the database, 4,790 individuals with GBC were identified with a prevalence rate of 8.5 per 100,000. Asian race has the highest prevalence of GBC (13.6/100,000). Patients with GBC were also more likely to be smokers, have a history of alcohol abuse, obesity, diabetes, cholelithiasis, chronic cholecystitis, primary sclerosing cholangitis (PSC), and chronic viral hepatitis.

Conclusions: This is one of the largest US population studies to date evaluating the epidemiology of GBC. The 20-year period prevalence rate of GBC was 8.5 per 100,000. Patients with GBC were more likely to be elderly, females, obese, diabetic, and have chronic hepatobiliary disorders.

Keywords: Epidemiology; gallbladder; neoplasm; risk factors

Submitted Oct 19, 2020. Accepted for publication Jan 17, 2021.
doi: 10.21037/cco-20-230
View this article at: http://dx.doi.org/10.21037/cco-20-230

Introduction

Gallbladder cancer (GBC) is a rare malignancy, ranking 22nd and 17th among cancers worldwide in incidence and mortality respectively (1). Nevertheless, it is the most common neoplasm of the biliary tract, accounting for 80–95% of all neoplasms of the biliary tract system (2). Most GBCs are adenocarcinomas that arise from the epithelial lining of the gallbladder and its carcinogenesis follows a progression through a metaplasia-dysplasia-carcinoma sequence. It is an aggressive tumor that usually presents late with metastasis; hence carrying a poor prognosis with a mean overall survival of 6 months after the diagnosis (2,3). In current practice, there is a lack of screening tests for GBC (4).

Current literature shows that unlike other gastroenterological
malignancies, GBC is more common in women compared to men (5). It is also more common in older individuals with 72 years old being the average age of diagnosis in the US (6). It is more prevalent in certain ethnicities and regions of the world compared to others which is attributed to different environmental exposures and genetic predisposition (5,7). Interestingly, the global distribution of GBC and gallbladder stones—the most commonly cited risk factor—is found to be related (5). Other risk factors reported in the literature include a family history of GBC (8), gallbladder polyps (9), primary sclerosis cholangitis (10), chronic infections with *Salmonella* (11) and *H. pylori* (12), congenital biliary cysts (13), abnormal pancreaticobiliary junction (14), hormonal therapy in postmenopausal women (15), cigarette smoking (16), heavy alcohol drinking (16,17), and obesity (18).

There are very few population-based studies regarding this topic (3,19-22). Using a large population database, we aim to describe the epidemiology and risk factors of GBC in the US.

We present the following article in accordance with the STROBE reporting checklist (available at http://dx.doi.org/10.21037/cco-20-230).

**Methods**

**Database**

This is a retrospective, cohort analysis using a multiple health system research platform (Explorys Inc., Cleveland, OH, USA) developed by IBM Corporation, Watson Health (23). Explorys include electronic health record (EHR) data of more than 50 million unique patients, from 26 healthcare systems with a total of 360 hospitals in the US from 1999 to date, representing approximately 15% of the population. Participating health systems are from all states and that provides a broad regional distribution of source population. Diagnoses, findings, and procedures are arranged into the Systematized Nomenclature of Medicine-Clinical Terms (SNOMED-CT) hierarchy (24) while prescription drug orders are mapped into SNOMED and RxNorm (25). Access to Explorys is granted to participating healthcare systems. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). Source data are de-identified; therefore, Institutional Review Board (IRB) is not required. To keep patients’ information confidentiality, Explorys rounds cell counts to the nearest 10. Explorys database has been validated previously in multiple fields including surgery and gastroenterology (26-36).

**Patient selection**

Using the Explorys platform, we identified a cohort of patients diagnosed with GBC between the period of December 1999 and December 2019. The study cohort (GBC) was identified by searching the database for a SNOMED-CT diagnosis of “primary malignant neoplasm of gallbladder” after excluding patients younger than 18 years old. The control group was then identified for those who have no GBC. Subsequently, a cohort of patients with first ever GBC diagnosis was identified between the period of December 2018 to December 2019 to calculate the incidence of GBC in the last year.

**Risk factors, and predisposing medical conditions associated with GBC**

Sex, age, and race-based data were collected. Possible risk factors and predisposing medical conditions included smoking, alcohol abuse, obesity, diabetes mellitus (DM), family history of gastrointestinal (GI) cancer, primary sclerosing cholangitis (PSC), chronic viral hepatitis [hepatitis B virus (HBV) and hepatitis C virus (HCV)], chronic cholecystitis, and cholelithiasis was investigated using SNOMED-CT diagnostic codes.

**Statistical analysis**

Demographics and associated diseases were characterized by descriptive statistics. The overall period prevalence was calculated by dividing the total number of individuals with GBC by the total number of individuals in Explorys [1999–2019], thus making sure that all patients in the denominator had an equal opportunity of being diagnosed with GBC. We also calculated the rate of new cases of GBC in the US in the past year [2018–2019] by dividing the number of individuals with first ever GBC diagnosis from December 2018 through December 2019 by the total number of individuals in Explorys in the past year in an attempt to provide a proxy estimate of annual incidence. A two-sided P value <0.05 was considered statistically significant, and all statistical analyses were performed using IBM SPSS Statistics version 25.

**Results**

**Descriptive epidemiology**

Baseline characteristics of the study cohort (GBC cohort)
and control are displayed in Table 1. Of the 56,197, 690 individuals in the database above the age of 18 years old [1999–2019], 4,790 individuals with GBC were identified. The 20-year period prevalence rate of GBC was 8.5 per 100,000. The incidence of new cases reported last year [2018–2019] was 7.4 per 100,000 population. Elderly patients above the age of 65 years old accounted for 77% of the cases, the highest prevalence was among the age group 80–90 years old (Figure 1). The period prevalence rate was higher among females compared to males (9.7 vs. 7.0 per 100,000) and higher in Asian (13.6 per 100,000) compared to Caucasian, African-American, and Hispanic (10.5, 11.4, and 6.6 per 100,000 respectively) (Figure 2).

Medical comorbidities associated with GBC patients

Patients with GBC had a higher prevalence of multiple comorbidities compared to the general population including smoking and alcohol abuse. There was significantly higher prevalence of biliary diseases such as cholelithiasis (31.3% vs. 1.6%), choledocholithiasis (34.4% vs. 1.7%), chronic cholecystitis (5.2% vs. 0.2%), choledochal cyst (0.1% vs. 0.001%), and PSC (0.1% vs. 0.01%). History of chronic viral hepatitis infections (HBV and HCV), obesity, and DM were also more prevalent in patients with GBC compared to the general population.

Risk factors for development of GBC

After adjusting using multivariate analysis model, patients with GBC were more likely to be elderly (above 65 years old) (OR: 6.12; 95% CI: 5.75–6.52) and females (1.35; 1.28–1.43). The risk of GBC varied among different race groups; compared to Caucasian, the risk of GBC among African-American, Asian, and Hispanic were (1.08; 1.01–1.17), (1.32; 1.11–1.56), and (0.77; 0.61–0.97) respectively. Patients who developed GBC had a history of smoking (1.73; 1.61–1.87) and alcohol abuse (1.69; 1.44–1.99). In terms of predisposing medical conditions, individuals with GBC were more likely to have a diagnosis of cholelithiasis (15.23; 14.27–16.27), chronic cholecystitis (4.20; 3.74–4.72), PSC (7.22; 5.01–10.41), obesity (3.30; 3.04–3.58), DM (2.13; 2.00–2.26), family history (4.80; 3.99–5.75), chronic HBV (3.79; 1.89–7.60), and chronic HCV (4.17; 3.15–5.51) (Figure 3).

Discussion

Our population-based study of 56 million patients showed that the 20-year period prevalence rate of GBC is 8.5 per 100,000 among the US population. In our cohort, patients with GBC were more likely to be elderly (age >65) and females.

Our study was consistent with previous epidemiologic studies in the US that showed a higher prevalence of GBC in women compared to men (22,37). GBC is considered one of the few cancers where the incidence is higher in women than men (5), this is thought to be due to a higher prevalence of gallstones in women (38). After adjusting for several factors in the multivariate analysis, we found that female gender is an independent risk factor for GBC. In our cohort, gallstone was the single most important risk for GBC, which is consistent with prior studies (5). Moreover, it is believed that the geographic pattern of GBC correlates with the prevalence of gallstones (7). Gallstone size is believed to influence the risk of GBC with larger stones (>3 cm) carrying a higher risk compared to smaller stones (39). However, despite the increased risk, the development of GBC in patients with gallstone disease remains rare (40–42). It is proposed that cholelithiasis leads to GBC through chronic inflammation and metaplasia (40). Chronic inflammation can also lead to the deposition of calcium in the gallbladder wall, a condition called porcelain gallbladder. Historically, porcelain gallbladder was an absolute indication for cholecystectomy due to the risk of GBC. However, today, the association between GBC and porcelain gallbladder is controversial (43,44).

The data on obesity as a risk factor for GBC is conflicting. Our study however, is in agreement with the previous studies that showed an increased risk of GBC in obese patients including a cohort of 181,000 individuals by Larsson et al. (45). In a systematic review of eight studies, obesity and body mass index (BMI) were significantly associated with GBC. The increased risk is thought to be related to increased predisposition to form gallstones. On the contrary, other studies have showed no association between obesity and GBC (46). Previous studies found an association between diabetes and GBC, but it was unclear whether this association was independent or mediated by obesity (47). We found that DM is independently associated with increased risk of GBC. Similarly, a recent European cohort study showed that diabetes was associated with a higher risk of GBC independent of obesity (48). Cigarette smoking and heavy alcohol use were also associated with GBC. In a prospective study conducted in Japan, there was a positive linear association observed between the risk of GBC and the number of cigarettes
Table 1 Baseline characteristics of patients with GBC and control group

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>GBC, N=4,790 (%)</th>
<th>Non-GBC, n=56,192,900 (%)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30–34</td>
<td>10 (0.2)</td>
<td>4,386,880 (7.8)</td>
<td></td>
</tr>
<tr>
<td>35–39</td>
<td>40 (0.8)</td>
<td>4,430,270 (7.9)</td>
<td></td>
</tr>
<tr>
<td>40–44</td>
<td>50 (1.0)</td>
<td>4,100,880 (7.3)</td>
<td></td>
</tr>
<tr>
<td>45–49</td>
<td>100 (2.1)</td>
<td>4,135,760 (7.4)</td>
<td></td>
</tr>
<tr>
<td>50–54</td>
<td>190 (4.0)</td>
<td>4,165,770 (7.4)</td>
<td></td>
</tr>
<tr>
<td>55–59</td>
<td>300 (6.3)</td>
<td>4,527,510 (8.1)</td>
<td></td>
</tr>
<tr>
<td>60–64</td>
<td>410 (8.6)</td>
<td>4,434,540 (7.9)</td>
<td>0.0690</td>
</tr>
<tr>
<td>65–69</td>
<td>590 (12.3)</td>
<td>4,034,890 (7.2)</td>
<td></td>
</tr>
<tr>
<td>70–74</td>
<td>730 (15.2)</td>
<td>3,542,690 (6.3)</td>
<td></td>
</tr>
<tr>
<td>75–79</td>
<td>710 (14.8)</td>
<td>2,800,150 (5.0)</td>
<td></td>
</tr>
<tr>
<td>80–84</td>
<td>630 (13.2)</td>
<td>2,230,810 (4.0)</td>
<td></td>
</tr>
<tr>
<td>85–90</td>
<td>520 (10.9)</td>
<td>1,832,690 (3.3)</td>
<td></td>
</tr>
<tr>
<td>90+</td>
<td>480 (10.0)</td>
<td>2,359,740 (4.2)</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Females</td>
<td>3,020 (63.0)</td>
<td>31,012,680 (55.2)</td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>3,420 (71.4)</td>
<td>32,430,430 (57.7)</td>
<td></td>
</tr>
<tr>
<td>African-American</td>
<td>650 (13.6)</td>
<td>5,683,420 (10.1)</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>130 (2.7)</td>
<td>953,840 (1.7)</td>
<td></td>
</tr>
<tr>
<td>Comorbidities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol abuse</td>
<td>130 (2.7)</td>
<td>915,620 (1.6)</td>
<td></td>
</tr>
<tr>
<td>Tobacco smoking</td>
<td>700 (14.6)</td>
<td>5,290,200 (9.4)</td>
<td></td>
</tr>
<tr>
<td>Drug abuse</td>
<td>210 (4.4)</td>
<td>1,754,020 (3.1)</td>
<td></td>
</tr>
<tr>
<td>Family history of GI cancer</td>
<td>240 (5.0)</td>
<td>500,750 (0.9)</td>
<td></td>
</tr>
<tr>
<td>Primary sclerosing cholangitis</td>
<td>5 (0.1)</td>
<td>8,010 (0.01)</td>
<td></td>
</tr>
<tr>
<td>Obesity</td>
<td>490 (10.2)</td>
<td>2,155,650 (3.8)</td>
<td></td>
</tr>
<tr>
<td>Metabolic syndrome x</td>
<td>40 (0.8)</td>
<td>184,920 (0.3)</td>
<td></td>
</tr>
<tr>
<td>DM</td>
<td>1,510 (31.5)</td>
<td>4,806,690 (8.6)</td>
<td></td>
</tr>
<tr>
<td>Chronic HBV</td>
<td>20 (0.4)</td>
<td>35,970 (0.1)</td>
<td></td>
</tr>
<tr>
<td>Chronic HCV</td>
<td>90 (1.9)</td>
<td>209,340 (0.4)</td>
<td></td>
</tr>
<tr>
<td>Liver cirrhosis</td>
<td>270 (5.6)</td>
<td>292,880 (0.5)</td>
<td></td>
</tr>
<tr>
<td>Primary liver cancer</td>
<td>2,100 (43.8)</td>
<td>169,890 (0.3)</td>
<td></td>
</tr>
<tr>
<td>Primary colon cancer</td>
<td>240 (5.0)</td>
<td>200,750 (0.4)</td>
<td></td>
</tr>
<tr>
<td>Primary pancreatic cancer</td>
<td>370 (7.7)</td>
<td>54,530 (0.1)</td>
<td></td>
</tr>
</tbody>
</table>

Table 1 (continued)
Table 1 (continued)

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>GBC, N=4,790 (%)</th>
<th>Non-GBC, n=56,192,900 (%)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liver fluke infection</td>
<td>5 (0.1)</td>
<td>130 (0.0002)</td>
<td></td>
</tr>
<tr>
<td>Choledochal cyst</td>
<td>5 (0.1)</td>
<td>610 (0.001)</td>
<td></td>
</tr>
<tr>
<td>Choledocholithiasis</td>
<td>1,650 (34.4)</td>
<td>973,500 (1.7)</td>
<td></td>
</tr>
<tr>
<td>Chronic cholecystitis</td>
<td>250 (5.2)</td>
<td>112,610 (0.2)</td>
<td></td>
</tr>
<tr>
<td>Choledolithiasis</td>
<td>1,500 (31.3)</td>
<td>914,820 (1.6)</td>
<td></td>
</tr>
<tr>
<td>Acute pancreatitis</td>
<td>390 (8.1)</td>
<td>325,190 (0.6)</td>
<td></td>
</tr>
<tr>
<td>Chronic pancreatitis</td>
<td>70 (1.5)</td>
<td>86,290 (0.2)</td>
<td></td>
</tr>
</tbody>
</table>

All P values are <0.0001 unless stated otherwise. GBC, gallbladder cancer; GI, gastrointestinal; DM, diabetes mellitus; HBV, hepatitis B virus; HBC, hepatitis C virus.

Figure 1 Prevalence rate of GBC among different age groups, per 100,000. GBC, gallbladder cancer.

smoked. The study showed a linear association between alcohol drinking and risk of GBC among men as well but not in women (16). In a meta-analysis investigating the association of alcohol and risk of 23 types of malignancy, the relative risk of GBC among heavy drinkers was 2.64 (95% CI: 1.62–4.30) (17).

Emerging data suggest that chronic HBV and HCV infections may also play a role in extrahepatic bile duct cancers. One Taiwanese study found that HBV and HCV infections were associated with increased risk of several extrahepatic malignancies including the risk of gallbladder and extrahepatic bile duct cancer compared to the general population (49). Similarly, a Chinese population-based study concluded that patients with chronic HBV have increased risk of extrahepatic bile duct cancer (50). In our study, we found a higher risk of developing GBC in HBV or HCV infection. Moreover, PSC was significantly associated with higher risk of GBC with a seven-fold increased risk compared to the general population. It is well-known that PSC is associated with an increased risk of several neoplasms, mainly cholangiocarcinoma, GBC, liver cancer, and colorectal neoplasm, with nearly 50% of deaths in...
patients with PSC being due to cancer (51).

The indolent and non-specific presentation of GBC along with a paucity of early screening methods, and predictive radiological features often preclude early detection of GBC. Therefore, that results in difficulty in implementing curative intent surgical resection and subsequently reflects negatively on the overall prognosis. GBC remains a highly fatal cancer, with only 10% of cases presenting at a surgically resectable stage with an overall 5-year survival of 60% for localized tumors, 30% for regional spread, and 2% for distant spread (52). Despite the importance of obtaining accurate prognostic information in helping physicians to make better clinical decisions, there are limited prognostic data in GBC at the current time. In a Chinese study, Bai et al. proposed a nomogram to predict the overall survival after GBC resection (53). The nomogram suggested that the absence of jaundice, lower preoperative CA19-9 levels, lower TNM stage, and incisional margins without tumor cells correlated well with a long survival time. Using the SEER database, Wang et al. formulated two nomograms that would help predict survival for GBC receiving chemoradiotherapy (54,55). On multivariate regression analysis, the model showed that age, sex, papillary histology, stage, and adjuvant radiotherapy

![Figure 2](image1.png)

**Figure 2** Prevalence rate of GBC based on sex and race. GBC, gallbladder cancer.

![Figure 3](image2.png)

**Figure 3** Risk factors of GBC. GBC, gallbladder cancer.
were significant predictors of overall survival. In a systemic review evaluating the prognostic value of DM on the survival of patients with GBC, a meta-analysis of 10 studies concluded that diabetic patients had a higher mortality of GBC (56). Extrapolating data from other cancers, smoking and alcohol abuse have detrimental effect post-cancer diagnosis as well as increased mortality (57,58).

This study has few limitations, mostly related to the use of a large database. Accurate prevalence rates of diagnoses might be affected as the database is generated from those who sought medical care. Given the de-identified nature of Explorys database, we were not able to obtain individual data regarding imaging reports and laboratory values, therefore we were not able to include patients with porcelain gallbladder or pathological data on gallbladder biopsies. Despite these limitations, the major strengths of this study lie in that it is one of the largest population-based studies to ever address GBC epidemiology and risk factors to our best knowledge. Explorys captures electronic medical record data of more than 50 million patients from all different regions of the US. In addition, SNOMED-CT allows for more concepts to be coded per clinical document compared to ICD making it more accurate in terms of documenting diagnoses and pertinent patient information. It is important to highlight that Explorys has been validated previously as a database (59) and in different specialties including gastroenterology (26-36,60,61).

**Conclusions**

Our study is in agreement with the previous literature in regards to the association of GBC with female sex, older age, gallstone disease, cholecystitis, obesity, diabetes, smoking, alcohol abuse, chronic HBV and HCV infections, family history of GBC, and PSC. This is the largest US study to-date demonstrating incidence rate higher than previously reported. GBC is a rare disease with a poor prognosis. GBC epidemiology is an evolving and is constantly reshaped by several factors including lifestyle, geographic, and diet. The clinician must keep a high index of suspicion of GBC in patients with the risk factors emphasized in this study for timely detection, diagnosis, and better outcomes.

**Acknowledgments**

*Funding:* None.

---

**Footnote**

*Reporting Checklist:* The authors have completed the STROBE reporting checklist. Available at [http://dx.doi.org/10.21037/cco-20-230](http://dx.doi.org/10.21037/cco-20-230)

*Data Sharing Statement:* Available at [http://dx.doi.org/10.21037/cco-20-230](http://dx.doi.org/10.21037/cco-20-230)

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at [http://dx.doi.org/10.21037/cco-20-230](http://dx.doi.org/10.21037/cco-20-230)). The authors have no other conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study did not require an IRB as data is de-identified.

*Open Access Statement:* This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: [https://creativecommons.org/licenses/by-nc-nd/4.0/](https://creativecommons.org/licenses/by-nc-nd/4.0/).

**References**


