

Digestive and Liver Diseases Statistics, 2004

MARK W. RUSSO, JEFFREY T. WEI, MICHELLE T. THINY, LISA M. GANGAROSA, ALPHONSO BROWN, YEHUDA RINGEL, NICHOLAS J. SHAHEEN, and ROBERT S. SANDLER

Division of Gastroenterology and Hepatology, Department of Medicine, and the Center for Gastrointestinal Biology and Disease, University of North Carolina, Chapel Hill, North Carolina

Background & Aims: Digestive and liver diseases are associated with substantial morbidity and mortality in the United States. Statistics about the incidence, prevalence, mortality, and resource utilization of digestive and liver diseases in the United States may be cumbersome to obtain because they are scattered in multiple sources. These data may be useful for policy makers, grant applicants, and authors. **Methods:** Data on the most common gastrointestinal and liver diseases were collected from large publicly available national databases. Information was collected on inpatient and outpatient gastrointestinal complaints and diagnoses, gastrointestinal cancers, and deaths from common liver diseases. **Results:** The leading gastrointestinal complaint prompting an outpatient visit is abdominal pain, with 12.2 million annual visits, followed by diarrhea, nausea, and vomiting. Abdominal pain is the leading outpatient gastrointestinal diagnosis, accounting for 5.2 million visits annually, followed by gastroesophageal reflux disease, with 4.5 million visits. Gallstone disease is the most common inpatient diagnosis, with 262,411 hospitalizations and a median inpatient charge of \$11,584. Colorectal cancer is the most common gastrointestinal cause of death and is the most common gastrointestinal cancer, with an incidence of 54 per 100,000. Among gastrointestinal cancers, primary liver cancer had the highest increase in incidence from 1992 to 2000. **Conclusions:** Gastrointestinal and liver diseases are associated with significant outpatient and inpatient healthcare utilization. Following trends in utilization is important for determining allocation of resources for health care and research.

Digestive and liver diseases are common in the United States. They affect large numbers of individuals and exact a considerable financial and social burden. These diseases are associated with frequent physician visits and hospitalizations and are sometimes fatal.

Although data on digestive and liver conditions are collected on a regular basis, they are sometimes difficult to locate. Moreover, previous compilations of statistics on digestive diseases are lengthy or do not provide data on outpatients.^{1,2} The American Gastroenterological Association

publication, *The Burden of Gastrointestinal Diseases*, is a comprehensive report, but it is lengthy and has not been recently updated.¹ Herein, we provide a concise description of frequently used statistics displayed in tables comparing common gastrointestinal diagnoses. A publication on selected digestive diseases in the United States focused mostly on costs and does not provide data on symptoms and diagnoses for outpatient visits or on the incidence and prevalence of gastrointestinal cancers.²

Our goals were to collect and report the most recent statistics on deaths, symptoms, physician visits, and hospitalizations for common digestive and liver diseases. These data are unique because they are population based and are gathered from large national databases. This information should be a valuable reference for authors, grant applicants, funding agencies, and policy makers. Importantly, the data are likely to raise questions and lead to future research.

Materials and Methods

Gastrointestinal Causes of Death

Data on the leading causes of death were obtained from mortality tables from the Division of Vital Statistics, National Center for Health Statistics (Table 1). A total of 2,403,351 deaths in the United States were reported during the year 2000. Causes of death were reported on the basis of the 10th revision of the International Classification of Diseases (ICD-10; 1992). In constructing these tables, the underlying cause of death is ascertained from death certificate data and then coded according to the most recent revision of ICD-10 codes. ICD-10 codes were published by the World Health Organization in 1992 and have been used for coding and classifying causes of death since 1999. ICD-10 codes differ from ICD-9 codes in some ways, although the overall content is similar. ICD-10 has alphanumeric categories rather than numeric categories and

Abbreviations used in this paper: ICD, International Classification of Diseases; NAMCS, National Ambulatory Medical Care Survey; NIS, Nationwide Inpatient Sample; SEER, Surveillance, Epidemiology, and End Results.

© 2004 by the American Gastroenterological Association
0016-5085/04/\$30.00

doi:10.1053/j.gastro.2004.01.025

Table 1. Leading 20 Gastrointestinal Causes of Death in the United States, 2000

Rank	Cause of death	No. of deaths ^a	ICD-10 codes
1	Colorectal/anal cancer	57,477	C18.0–C18.9, C19, C20, C21.0–C21.2, C21.8
2	Pancreatic cancer	29,332	C25.0–C25.4, C25.8–C25.9
3	Fibrosis/cirrhosis of liver and hepatic failure not otherwise specified	14,003	K72.0–K72.1, K72.9, K74.0–K74.1, K74.3–K74.6
4	Malignant neoplasms of the liver and intrahepatic ducts	12,916	C22.0–C22.4, C22.7, C22.9
5	Stomach cancer	12,645	C16.0–C16.9
6	Esophageal cancer	12,232	C15.0–C15.9
7	Alcoholic liver disease	12,109	K70.0–K70.4, K70.9
8	Vascular disorders of the intestine	9060	K55.0–K55.2, K55.8–K55.9
9	Gastrointestinal hemorrhage, unspecified	7719	K92.2
10	Paralytic ileus and intestinal obstruction	5221	K56.0–K56.7
11	Ulcers (gastric/duodenal/peptic)	4504	K25.0–K25.1, K25.3–K25.7, K25.9, K26.0–K26.7, K26.9, K27.0–K27.7, K27.9, K28.4–K28.9
12	Acute hepatitis C	3702	B17.1
13	Diverticular disease	3290	K57.0–K57.5, K57.8–K57.9
14	Acute pancreatitis	2834	K85
15	Malignant neoplasms of the gallbladder	1949	C23
16	Biliary tract cancer	1717	C24.0–C24.1, C24.8–C24.9
17	Peritonitis	1661	K65.0, K65.8–K65.9
18	Perforation of intestine (nontraumatic)	1571	K63.1
19	Cholecystitis	1430	K81.0–K81.1, K81.8–K81.9
20	<i>Clostridium difficile</i> enterocolitis	1101	A04.7

NOTE. Source: National Center for Health Statistics Web site—mortality tables (<http://www.cdc.gov/nchs/data/dvs/wktbli.pdf>). Accessed: October 7, 2003.

^aOf a total of 2,403,351 deaths in 2000.

has almost twice as many categories as ICD-9. Therefore, data presented with ICD-9 codes cannot be directly compared with data presented as ICD-10 codes because of differences in the terms used for specific diagnoses. Of note, gastrointestinal malignancies have not been subclassified by histological subtype.

Gastrointestinal Symptoms and Diagnoses

The leading gastrointestinal symptoms prompting an outpatient clinic visit were estimated from a random sample of 27,369 patient visits to 1388 physicians participating in the National Ambulatory Medical Care Survey (NAMCS) 2000 (Tables 2 and 3). This survey was also used to estimate the leading physician diagnoses for outpatient clinic visits. The NAMCS was a national survey designed to provide information about the use of ambulatory care services in the United States. Findings were based on a sample of visits to non-federally employed office-based physicians who were primarily engaged in direct patient care. Each physician was randomly assigned to a 1-week reporting period. During that period, physicians, office staff, or both recorded data from a random sample of visits. Data included patients' symptoms and physicians' diagnoses. Patients were asked to provide, in their own words, their primary complaint, symptom, or other reason for the office visit. Patients provided their complaint, and after the office visit, physicians or staff were asked to code the primary diagnosis for the visit and the primary complaint. In 2000, a total of 27,369 patient record forms were received from the 1388 physicians who participated in the NAMCS. To obtain national estimates from the sample, an inflation factor called

Table 2. Leading Gastrointestinal Symptoms Prompting an Outpatient Clinic Visit in the United States, 2000

Symptom	Estimated no. of visits ^a
Abdominal pain, cramps, spasms	12,275,373
Diarrhea	4,061,491
Nausea	3,316,009
Vomiting	2,884,373
Heartburn and indigestion (dyspepsia)	1,824,043
Constipation	1,327,156
Anal/rectal bleeding	1,260,644
Blood in stool (melena)	1,179,224
Other, unspecified gastrointestinal symptoms	957,866
Decreased appetite	715,737
Difficulty swallowing (dysphagia)	676,421
Anal/rectal pain	514,896
Other symptoms or changes in bowel function	466,803
Jaundice	431,410
Abdominal distension, fullness, or swelling	384,877
Abdominal mass or tumor	247,850
Gastrointestinal bleeding	229,729
Anal/rectal swelling or mass	224,093
Flatulence	209,434

NOTE. Source: National Ambulatory Medical Care Survey 2000 (<http://www.cdc.gov/nchs/nhcs.htm>).

^aThe total number of outpatient visits was estimated from a random sample of 27,369 patient visits to 1388 physicians participating in the National Ambulatory Medical Care Survey 2000.

Table 3. Leading Physician Diagnoses for Gastrointestinal Disorders in Outpatient Clinic Visits in the United States, 2000^a

Rank	Diagnosis	ICD-9 codes	Estimated no. of visits
1	Abdominal pain	789.00–789.09	5,241,850
2	Gastroesophageal reflux disease	530.11, 530.81	4,591,403
3	Gastroenteritis	558.90	3,431,259
4	Gastritis	535.00, 535.40, 535.50	2,400,034
5	Hemorrhoids	455.00–455.90	1,568,274
6	Irritable bowel syndrome	564.10	1,557,133
7	Hernias—noninguinal	551.00–554.90	1,544,000
8	Benign neoplasm of colon	211.30	1,516,086
9	Colorectal cancer	153.90, 154.00, 154.10, V10.05	1,491,134
10	Inguinal hernia	550.00–550.93	1,242,747
11	Diverticulosis of colon	562.10, 562.11	1,072,545
12	Dyspepsia	536.80, 536.90	1,001,114
13	Constipation	564.00	794,047
14	Gallstones	574.00–574.50	778,632
15	Occult gastrointestinal blood loss	792.10	762,527
16	Acute or chronic hepatitis C	070.51, 070.54	756,774
17	Crohn's disease	555.90	725,157
18	Anal/rectal bleeding	569.30	596,948
19	Ulcerative colitis or proctitis	556.20, 556.90	487,893
20	Gastric or duodenal ulcer	531.00–533.90, V12.71	457,153
21	Abnormal liver scan	794.80	443,310
22	Cirrhosis	571.20, 571.50, 571.60	419,021
23	Melena	578.10	401,602

NOTE. The total number of outpatient visits was estimated from a random sample of 27,369 patient visits to 1388 physicians participating in the National Ambulatory Medical Care Survey 2000 (<http://www.cdc.gov/nchs/nhcs.htm>).

^aThe numbers of visits for the diagnoses below rank 10 were estimated from weighting a small number of office encounters (<30) and may be unreliable.

patient visit weight was applied to estimate statistics for the 823,541,999 office visits made in the United States during the year 2000. Table 2 describes the leading patient complaints, and Table 3 lists the subsequent physician diagnosis.

Hospitalizations for Gastrointestinal Disorders

Data on hospitalizations for gastrointestinal diseases were obtained from the Healthcare Costs Utilization Project

(Table 4). This project is composed of the Nationwide Inpatient Sample (NIS), which is a unique and powerful database of hospital inpatient stays. NIS contains all discharge data from 994 hospitals located in 28 States, approximating a 20% stratified sample of US community hospitals with data from approximately 7.5 million hospital stays. NIS 2000 is a sample of hospitals that comprise approximately 80% of all hospital discharges in the United States. NIS is the only national hospital database with charge information on all patients,

Table 4. Most Common Inpatient Principal Gastrointestinal Diagnoses in the United States, 2000

Rank among gastrointestinal diagnoses	Rank among all diagnoses	Diagnosis (ICD-9 code)	Total no. of discharges	Median length of stay (days)	Median charges (\$)	Deaths (%)
1	574 #35 ^a 574.1 #43 ^b	Cholelithiasis with acute cholecystitis (574); cholelithiasis with cholecystitis not elsewhere classifiable (574.1)	262,411	2	11,584	0.6
2	20	Acute pancreatitis (577)	210,188	4	9364	1.8
3	27	Acute appendicitis (540.9)	170,337	2	8525	— ^c
4	32	Diverticulitis of colon without hemorrhage (562.11)	155,407	4	9468	1.1
5	34	Noninfectious gastroenteritis (574.1)	147,939	2	3990	0.3
6	45	Gastrointestinal hemorrhage (578.9)	115,355	3	7751	5
7	56	Esophageal reflux (530.81)	96,019	2	6712	— ^c
8	61	Intestinal obstruction (560.9)	93,226	3	6479	3.1
9	84	Diverticulosis with hemorrhage (562.12)	68,292	3	7804	1.2

NOTE. Source: Healthcare Costs Utilization Project (<http://hcup.ahrq.gov/HCUPNet.asp>).

^aThe individual ICD-9 code for cholelithiasis with acute cholecystitis (574) and its rank among all diagnoses (#35).

^bThe individual ICD-9 code for cholelithiasis with cholecystitis not elsewhere classifiable (574.1) and its rank among diagnoses (#43).

^cThe number of deaths was too small for a stable estimate.

Table 5. Number of Reported Cases of Infectious Gastrointestinal Diseases in the United States, 2002

Disease	n
Salmonellosis	11,056
Giardiasis	6660
Shigellosis	5127
Acute hepatitis A	3767
Acute hepatitis B	2767
Enterohemorrhagic <i>Escherichia coli</i>	527

NOTE. Source: *Morbidity and Mortality Weekly Report* Table II (part 2, 3, 4, and 6), week ending May 17, 2003 (MMWR Morb Mortal Wkly Rep 2003;52:478–486 (<http://www.cdc.gov/mmwr/>)).

regardless of payer, including persons covered by Medicare, Medicaid, private insurance and the uninsured. In Table 4, the leading gastrointestinal diagnoses (primary diagnosis) for hospitalizations are provided. The diagnosis of cholecystitis was combined from 2 ICD-9 codes (cholelithiasis with acute cholecystitis [574] and cholelithiasis with cholecystitis [574.1]) to accurately reflect the burden of disease from gallstones, because these diagnoses are very similar. The individual rankings are provided in the table.

Infectious Gastrointestinal Diseases and Liver Diseases

Information on cases of infectious gastrointestinal diseases was obtained from the *Morbidity and Mortality Weekly Report* (Table 5). The 50 states, the District of Columbia, and US territories report data on reportable diseases. Diseases that are reportable and of interest to gastroenterologists are listed in Table 5. Data on reportable diseases are updated on a regular basis with comparison made to the previous year. The data are revised weekly and are subject to change. As additional information becomes available, statistics are updated and revised.

Mortality From Liver and Biliary Diseases

Table 6 lists the mortality burden from the leading liver diseases. Causes of death from selected liver diseases were obtained from the United States Department of Health and Human Services, the Centers for Disease Control and Prevention, the National Center for Health Statistics, the Office of Analysis and Epidemiology, the Compressed Mortality File, computer tape, and the Centers for Disease Control and Prevention WONDER online database. These sources are public records with links to sites that collect population-based data on human diseases. The Compressed Mortality File is a county-level national mortality and population database that spans decades; it is derived from US records of deaths (death certificates) since 1979. Counts and rates of death can be obtained by place of residence (US state and county), age (17 age groups), race (white, black, and other), sex, year, and underlying cause of death (4-digit ICD code or group of codes). Death rates based on counts of 20 or less (death count ≤ 20) are flagged as unreliable. A death rate based on fewer than 20

deaths has a relative SE of $\geq 23\%$, which is considered statistically unreliable.

Gastrointestinal Cancer Rates

Information on cancer rates is based on data published by the Surveillance, Epidemiology, and End Results (SEER) program (<http://seer.cancer.gov>) of the National Cancer Institute (<http://www.cancer.gov>) (Table 7). The SEER program contains population-based information on cancer incidence and survival in the United States. It provides data collected and published, from 1975, on cancer prevalence, incidence, and mortality rates from 9 population-based cancer registries including the San Francisco/Oakland standard metropolitan statistical area, Connecticut, Detroit (metropolitan), Hawaii, Iowa, New Mexico, Seattle (Puget sound), Utah, and Atlanta (metropolitan). The SEER program uses a “first malignant primary only” method. The data include only the first malignant tumor and count a person only once (i.e., the prevalence of specific cancer sites adds up to the prevalence of all cancers combined). The reported prevalence rates for gastrointestinal cancers (Table 7) include all alive patients who were diagnosed with gastrointestinal cancer before 2000. Once a case is counted, it will be included in the prevalence of the specific diagnosis as long as the individual is alive, even if the individual was treated and cured (e.g., the patient underwent an operation and the cancer was excised). In addition, the SEER database counts only the first diagnosed primary cancer. Thus, an individual case can be included only once and would not be counted again if a second primary cancer were diagnosed.

Discussion

The data presented in this article provide a comprehensive and contemporary summary of digestive and liver disease in the United States. By compiling data from disparate sources, we provide a convenient resource

Table 6. Causes of Death From Selected Liver Diseases in the United States, 1999

Disease	Rate per 100,000	No. of deaths	ICD-10 code
Liver cell carcinoma	4.5	12,383	C22
Alcoholic liver disease	4.4	11,958	K70.0–K70.4, K70.9
Acute hepatitis C	1.2	3181	B17.1
Bile duct carcinoma	0.9	2552	C22.1
Acute hepatitis B	0.3	703	B16, B16.0, B16.1, B16.2, B17.0
Chronic hepatitis C	0.2	582	B18.2
Primary biliary cirrhosis	0.2	414	K74.3
Acute hepatitis A	— ^a	134	B15, B15.0, B15.9
Chronic hepatitis B	— ^a	129	B18, B18.1

NOTE. Source: United States Department of Health and Human Services, Centers for Disease Control and Prevention (CDC), National Center for Health Statistics, Office of Analysis and Epidemiology, Compressed Mortality File, computer tape, and CDC WONDER online database.

^aThere were too few cases to calculate a stable estimate.

Table 7. Prevalence and Incidence of Gastrointestinal Cancers in the United States, 1996–2000

Rank among gastrointestinal cancers	Rank among all sites	Cancer site	Prevalent cases ^a	Incidence ^a	Incidence trend ^b	Mortality ^c	Mortality trend ^b
1	3	Colon and rectum	1,041,499	54.2	−0.6	21.2	−1.7 ^d
2	19	Stomach	59,527	9.1	−1.3 ^d	4.8	−2.8 ^d
3	21	Pancreas	23,534	11.1	−0.4	10.5	−0.1
4	24	Esophagus	21,528	4.5	0.3	4.3	0.6 ^d
5	25	Liver and intrahepatic ducts	13,266	5.9	3.9 ^d	4.5	2.1 ^d
6	NA	Gallbladder	5749	1.2	−2.6 ^d	0.8	−2.4 ^d

NOTE. Source: Surveillance, Epidemiology, and End Results (SEER) program (www.seer.cancer.gov).

NA, statistics not available.

^aThe reported prevalence rates represent the estimated US cancer prevalence counts based on the 2000 cancer prevalence from the SEER registries and January 1, 2000, US population estimates. These prevalences are for all ages, races/ethnicities, sexes, and diseases diagnosed. The reported incidence rates represent the age-adjusted incidence per 100,000 for all ages, races/ethnicities, sexes, and diseases diagnosed between 1996 and 2000, from the 12 SEER registries.

^bTrends in SEER incidence and US mortality are reported by annual percentage change for 1992–2000 and for all ages, races/ethnicities, and sexes. The mortality trends are per 100,000 and are age-adjusted to the 2000 US standard population.

^cThe reported mortality represents the US death rates per 100,000, based on data provided by the National Center for Health Statistics public use data file (<http://www.cdc.gov/nchs>). The data are for the entire US population between 1996 and 2000 and are age-adjusted to the 2000 US standard population.

^dThe annual percentage change is significantly different from zero ($P < 0.05$).

for researchers and health planners. In addition, statistics can be periodically updated to identify trends in incidence, prevalence, and mortality of diseases, and this may have implications for health policy.

A strength of the data assembled for this report is that the data were population based and can be used to calculate national statistics. In addition, the NAMCS relies on trained interviewers who visited physicians before participation to instruct them how to complete the forms. All medical coding and keying operations were subject to quality-control procedures.

The databases that we used in the report have their limitations. For example, we used NAMCS to derive data on symptoms and office diagnoses. In the NAMCS, only visits to non–federally employed physicians are included, and therefore data do not apply to veterans and other populations that receive care within a self-contained system. Encounters outside of the physician's office, including telephone contact, house calls, hospital settings, and institutional settings, are not included. In addition, estimates based on fewer than 30 records should be considered unreliable. The NAMCS may overstate the outpatient burden due to colorectal cancer because patients may be seen regularly for surveillance after treatment compared with patients with gallstone disease, who may not be seen after cholecystectomy.

A limitation of mortality data is that trends in death rates are difficult to establish because changes in coding from using ICD-9 codes to ICD-10 codes occurred in 1999. Therefore, comparisons in mortality from before and after 1999 are difficult to perform. If comparisons are to be made over time, it is important to be consistent in

coding and classifying diseases from year to year. For example, we found that the number of deaths from alcoholic liver disease was much different from the number of deaths from alcoholic cirrhosis. Consistency in diagnosis and coding is necessary so that estimates accurately reflect trends that may occur over time. In addition, misclassification of diagnoses may occur. For instance, in 1999 and 2000 (Tables 1 and 6), more than 3000 individuals are reported to have died from acute hepatitis C. These estimates are higher than expected given that fatal acute hepatitis C is rarely reported in the literature. One possible explanation is that these patients may have been misclassified or misdiagnosed and actually died from chronic hepatitis C.

Rates of reportable infectious diseases are published in the *Morbidity and Mortality Weekly Report*. Data are derived from state and local health departments, and therefore the accuracy of these rates is dependent on physicians complying with reporting guidelines. Statistics on infectious gastrointestinal diseases may be underestimated if data on cases are missing or incomplete. In addition, the number of cases of infectious diseases may be underestimated if the disease is asymptomatic, as is the case with acute or chronic hepatitis C. Despite these limitations, the *Morbidity and Mortality Weekly Report* provides valuable information on the incidence of these diseases.

The SEER registry reports data on cancer prevalence. Patients diagnosed with cancer are considered prevalent cases until death, even if they had an early-stage cancer that was considered cured by resection. Therefore, prevalence, in this case, includes patients who are living with cancer (true prevalence) and patients who had cancer but

are now cured. For instance, the SEER registry captures data from 12 states, and statistics on the national cancer burden are extrapolated to the entire US population. Because the SEER registry collects data from selected sites, calculations based on SEER registry data may not truly reflect cancer rates in the United States because of differences in cancer rates in geographic areas not included in SEER.

The first step in decreasing the burden of digestive and liver diseases is to improve our understanding of the number of cases and the impact. It is to be hoped that the information included in this report will help to focus debate and research on the more common diseases and symptoms that bring patients to physicians. Success will be measured by decreases in incidence, prevalence, and cost when the data are next updated.

References

1. The burden of gastrointestinal diseases. Bethesda, MD: American Gastroenterological Association, 2001.
2. Sandler RS, Everhart JE, Donowitz M, Adams E, Cronin K, Goodman C, Gemmen E, Shah S, Avdic A, Rubin R. The burden of selected digestive diseases in the United States. *Gastroenterology* 2002;122:1500–1511.

Received July 29, 2003. Accepted January 21, 2004.

Address requests for reprints to: Robert S. Sandler, M.D., M.P.H., CB#7555, 4111 Bioinformatics Building, University of North Carolina, Chapel Hill, North Carolina 27599-7555. e-mail: rsandler@med.unc.edu; fax: (919) 966-6842.

This research was supported, in part, by grants P30DK34987 and T32DK07634 from the National Institutes of Health and by the American Gastroenterological Association Merck Clinical Research Career Development Award (to M.W.R.).