Estimation of State-Level Prevalence of Hepatitis C Virus Infection, US States and District of Columbia, 2010

Eli S. Rosenberg,1 Eric W. Hall,1 Patrick S. Sullivan,1 Travis H. Sanchez,1 Kimberly A. Workowski,2 John W. Ward,3 and Deborah Holtzman3

1Department of Epidemiology, Emory University Rollins School of Public Health; 2Division of Infectious Diseases, Department of Medicine, Emory University School of Medicine; and 3Division of Viral Hepatitis, National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, Centers for Disease Control and Prevention, Atlanta, Georgia

Background. Hepatitis C virus (HCV) infection is the most common chronic blood-borne infection in the United States and a leading cause of morbidity and mortality. Previous analyses of the US National Health and Nutrition Examination Survey (NHANES) indicated approximately 3.6 million noninstitutionalized persons with antibody to HCV (anti-HCV). However, state-level prevalence remains less understood and cannot be estimated reliably from NHANES alone.

Methods. We used 3 publicly available government data sources to estimate anti-HCV prevalence in each US state among noninstitutionalized persons aged ≥18 years. A small-area estimation model combined indirect standardization of NHANES-based prevalence with logistic regression modeling of mortality data, listing acute or chronic HCV infection as a cause of death, from the National Vital Statistics System during 1999–2012. Model results were combined with US Census population sizes to estimate total number and prevalence of persons with antibody to HCV in 2010.

Results. National anti-HCV prevalence was 1.67% (95% confidence interval [CI], 1.53–1.90), or 3 911 800 (95% CI, 3 589 400–4 447 500) adults in 2010. State-specific prevalence ranged from 0.71% (Illinois) to 3.34% (Oklahoma). The West census region had the highest region-specific prevalence (2.14% [95% CI, 1.96–2.48]); 10 of 13 states had rates above the national average. The South had the highest number of persons with anti-HCV (n = 1 561 600 [95% CI, 1 427 700–1 768 900]). The Midwest had the lowest region-specific prevalence (1.14% [95% CI, 1.04%–1.30%]).

Conclusions. States in the US West and South have been most impacted by hepatitis C. Estimates of HCV infection burden are essential to guide policy and programs to optimally prevent, detect, and cure infection.

Keywords. hepatitis C; surveillance; prevalence; National Health and Nutrition Examination Survey; National Vital Statistics System.

Hepatitis C virus (HCV) infection is the most common blood-borne infection in the United States [1]; untreated infection is a leading cause of cirrhosis and hepatocellular carcinoma, is the most frequent cause of liver failure requiring transplantation, and causes more deaths annually than human immunodeficiency virus (HIV) [2–5]. About 70% of persons with untreated HCV infection remain infected for life. HCV infection may be asymptomatic and approximately half of persons are unaware of their infection [6]. Within 30 years, 41% of infected persons progress to cirrhosis, leading to liver failure, hepatocellular carcinoma, and death from liver-related causes [2].

The annual number of new HCV infections in the United States was highest before HCV was discovered to be the cause of “non-A non-B hepatitis” in 1989; incidence declined after prevention guidelines and blood donor screening were implemented in the 1990s. After years of level incidence, new HCV infections began to rise nationally in 2010, increasing >2-fold by 2014. Incidence of acute HCV infections has risen among males, persons aged 20–29 years, and American Indians/Alaska Natives [7]. The prevalence of HCV infection can be measured by a positive test for HCV antibody. Chronic HCV infection is associated with substantial morbidity and mortality, is highest among males and those born during 1945–1965, and has been associated with rising HCV-related mortality in recent years [1, 4, 7–9].

Current US surveillance programs provide incomplete estimates of HCV infection prevalence. In 2014, the Centers For Disease Control and Prevention (CDC) National Notifiable Diseases Surveillance System (NNDSS) included passively collected laboratory reports from 40 states of acute HCV infection, and from 34 states reporting chronic HCV infection. US national and state-specific estimates of HCV infection prevalence cannot be ascertained from NNDSS data [7, 10].
National estimates of HCV prevalence have been produced using the US National Health and Nutrition Examination Survey (NHANES), which includes interviews and physical examinations of noninstitutionalized persons aged ≥26 years [11]. NHANES data for 2003–2010 estimated that 3.6 million persons had antibodies to HCV (anti-HCV), indicative of past or current HCV infection, corresponding to a national prevalence of 1.3% [8]. Subsequent systematic reviews estimate an additional 1 million HCV infections from populations excluded and underrepresented in NHANES, primarily homeless or incarcerated persons [12, 13].

State-level estimates of the prevalence of HCV infection are essential for guiding intervention programs, research, and federal assistance funding priorities among US states. In 2016, CDC funded 7 jurisdictions to conduct active case surveillance for HCV infection, and case counts of HCV diagnoses have been produced [7]. However, these states have been conducting active surveillance for few years, and most states lack local resources to provide similar information regarding the number of persons diagnosed with HCV; surveillance data are limited by variability in testing of persons at risk for infection and reporting to public health authorities.

Some jurisdictions have created their own estimates of prevalence using various methods, most commonly by applying NHANES-derived HCV infection prevalence rates to state populations [14]. The assumption that all states have the national NHANES prevalence of HCV infection might produce inaccurate state-specific estimates, because risk of HCV infection likely varies by state. Some researchers have attempted to improve upon this simple approach by standardizing the national NHANES estimate to local demographic profiles or creating models using case surveillance data (Supplementary Table 1). Despite these informative efforts, there is no complete set of state-specific estimates of HCV infection prevalence for all US states that is based on accurate and consistent methods.

Small-area estimation methods are statistical approaches that can be used for determining state-level HCV infection prevalence by allocating national prevalence estimates into state-specific components using data sources indicating state-level markers of HCV infection prevalence [15]. State-level data sources include electronic medical records (EMRs), insurance claims, and laboratory and mortality data. Individual commercially available EMR, claims, and laboratory datasets are inconsistent in their geographic and demographic representation [16]. Alternatively, mortality is systematically recorded for all decedents in the United States and is publicly available through the population-based National Vital Statistics System (NVSS) [4]. NVSS mortality data provide detailed information regarding decedent age, sex, and race/ethnicity, which are covariates of state-specific HCV infection prevalence.

Here, we describe an application of small-area methodology to deconstruct national NHANES-estimated HCV infection prevalence into state-specific components using HCV-related mortality as a state-level data source. Sensitivity

---

Table 1. Data Sources and Purposes for Primary and Sensitivity Analyses

| Data source | Years Represented | Purpose | No. of Individuals Represented | No. of Cases | Data Extraction Notes |
|-------------|-------------------|---------|-------------------------------|-------------|----------------------|
| National Health and Nutrition Examination Survey | 1999–2012 | National HCV antibody prevalence overall and by strata of sex, race/ethnicity, and birth cohort | 36,726 with nonmissing HCV antibody test results | 662 with positive anti-HCV test result | NHANES 2000, 2002, 2004, 2006, 2008, 2010, 2012 datasets |
| US Census inter-censal data | 1999–2012 | Population structure for modeling 14-year average HCV-related mortality rate | 3,125,647 person-years aged ≥18 y | NA | US Vintage 2000, Vintage 2009, Vintage 2014 datasets |
| US Census 2010 | 2010 | Population structure for final estimates | 234,564,071 persons aged ≥18 y | NA | US Census 2010 dataset |
| National Vital Statistics System | 1999–2012 | Hepatitis C-related mortality, for primary analysis | 33,540,118 decedents aged ≥18 y who resided in the 50 states or Washington, DC | 185,285 with HCV as underlying or multiple cause of death | ICD-10 codes included acute viral hepatitis C (B17.1), chronic viral hepatitis C (B18.2) |
| National Vital Statistics System | 1999–2012 | Cirrhosis-related mortality, for bias analysis | 33,540,118 decedents aged ≥18 y who resided in the 50 states or Washington, DC | 427,404 with cirrhosis as underlying or multiple cause of death | ICD-10 codes included hepatic fibrosis (K74.0); hepatic sclerosis (K74.1); hepatic fibrosis with hepatic sclerosis (K74.12); or other and unspecified cirrhosis of liver (K74.6) |
| National Vital Statistics System | 1999–2012 | Hepatocellular carcinoma-related mortality, for bias analysis | 33,540,118 decedents aged ≥18 y who resided in the 50 states or Washington, DC | 197,976 with hepatocellular carcinoma as underlying or multiple cause of death | ICD-10 codes included liver cell carcinoma (C22.0); primary malignant neoplasm of liver, type unspecified (C22.8); or malignant neoplasm of liver, not specified as primary or secondary (C22.9) |

Abbreviations: HCV, hepatitis C virus; ICD-10, International Classification of Diseases, Tenth Revision; NHANES, National Health and Nutrition Examination Survey.
analyses consider the uncertainty of the estimates associated with state-specific variation in diagnosing and treating HCV infection and in recording HCV-related mortality. We present state-level estimates of anti-HCV prevalence and compare these to previously published state estimates.

METHODS

We employed a synthetic small-area estimation approach that combined indirect standardization of NHANES data with regression-based estimates of state-level HCV-related mortality. The data sources and approach are described below and in Table 1 and Figure 1.

Data Sources

NHANES 1999–2012

Annually, NHANES uses a complex, multistage sampling design to select approximately 5000 persons in 15 counties, representing the noninstitutionalized population of persons aged ≥6 years residing in the 50 states and Washington, District of Columbia [11].

Nestor, with state-specific variation in diagnosing and treating HCV infection and in recording HCV-related mortality. We present state-level estimates of anti-HCV prevalence and compare these to previously published state estimates.

METHODS

We employed a synthetic small-area estimation approach that combined indirect standardization of NHANES data with regression-based estimates of state-level HCV-related mortality. The data sources and approach are described below and in Table 1 and Figure 1.

Data Sources

NHANES 1999–2012

Annually, NHANES uses a complex, multistage sampling design to select approximately 5000 persons in 15 counties, representing the noninstitutionalized population of persons aged ≥6 years residing in the 50 states and Washington, District of Columbia [11].

Nestor, with state-specific variation in diagnosing and treating HCV infection and in recording HCV-related mortality. We present state-level estimates of anti-HCV prevalence and compare these to previously published state estimates.

METHODS

We employed a synthetic small-area estimation approach that combined indirect standardization of NHANES data with regression-based estimates of state-level HCV-related mortality. The data sources and approach are described below and in Table 1 and Figure 1.

Data Sources

NHANES 1999–2012

Annually, NHANES uses a complex, multistage sampling design to select approximately 5000 persons in 15 counties, representing the noninstitutionalized population of persons aged ≥6 years residing in the 50 states and Washington, District of Columbia [11].

Nestor, with state-specific variation in diagnosing and treating HCV infection and in recording HCV-related mortality. We present state-level estimates of anti-HCV prevalence and compare these to previously published state estimates.

METHODS

We employed a synthetic small-area estimation approach that combined indirect standardization of NHANES data with regression-based estimates of state-level HCV-related mortality. The data sources and approach are described below and in Table 1 and Figure 1.

Data Sources

NHANES 1999–2012

Annually, NHANES uses a complex, multistage sampling design to select approximately 5000 persons in 15 counties, representing the noninstitutionalized population of persons aged ≥6 years residing in the 50 states and Washington, District of Columbia [11].

Nestor, with state-specific variation in diagnosing and treating HCV infection and in recording HCV-related mortality. We present state-level estimates of anti-HCV prevalence and compare these to previously published state estimates.

METHODS

We employed a synthetic small-area estimation approach that combined indirect standardization of NHANES data with regression-based estimates of state-level HCV-related mortality. The data sources and approach are described below and in Table 1 and Figure 1.

Data Sources

NHANES 1999–2012

Annually, NHANES uses a complex, multistage sampling design to select approximately 5000 persons in 15 counties, representing the noninstitutionalized population of persons aged ≥6 years residing in the 50 states and Washington, District of Columbia [11].

Nestor, with state-specific variation in diagnosing and treating HCV infection and in recording HCV-related mortality. We present state-level estimates of anti-HCV prevalence and compare these to previously published state estimates.
District of Columbia and were aged ≥18 years at death. Records also contained International Classification of Diseases, Tenth Revision (ICD-10) codes for multiple underlying causes of death (N = 33 540 118). Demographic covariates included sex, race/ethnicity, and birth cohort. In the primary analysis, any records including the ICD-10 code for acute viral hepatitis C (B17.1) or chronic viral hepatitis C (B18.2) as an underlying cause of death were considered to signal HCV-related mortality (n = 185 285).

**US Census Intercensal Data (1999–2012) and 2010 US Census Data**

Intercensal population estimates from the US Vintage 2000, 2009, and 2014 datasets provided denominators for HCV-related mortality rates during 1999–2012 [21]. Data were grouped into the same sex, race/ethnicity, and birth categories as the NHANES and NVSS data. State-by-demographic group population totals from the 2010 US Census were used to calculate 2010 state HCV prevalence [21].

**Analysis**

The number of persons in each state with anti-HCV was computed using the standardization-based estimator in Figure 1, equation 1. First, we calculated direct weighted estimates of national HCV-antibody prevalence \( \hat{\lambda}_i \) for 24 strata (sex × race/ethnicity × birth cohort), using standard methodology (Figure 1, equation 1) [15]. We multiplied weighted estimates by state-by-demographic stratum 2010 population counts to generate crude state-level estimates. These were adjusted by the ratio of state-by-demographic stratum effects, based on the average HCV-related death rate \( \theta_j \) in the 14-year period. We fit a high-order logistic regression model that approximated full stratification (several of 1224 strata had zero cells), permitting detection of heterogeneity among strata (Figure 1, equation 3). We assessed collinearity and model fit by comparing observed state-level HCV-related mortality totals to model predictions [15, 22]. Mortality-adjusted HCV infection prevalence totals were summarized to yield estimated state-level totals \( T_i \) (rounded to nearest hundred persons), with prevalence rates \( \hat{\lambda}_i \) (Figure 1, equations 1 and 2). Supplementary analyses estimated state-level chronic HCV infection, defined as a positive or indeterminate anti-HCV test and a positive HCV RNA test, using the above approach but with a 12-stratum model (race/ethnicity considered white non-Hispanic or not), due to more sparse NHANES data.

**Sensitivity Analyses**

To account for the joint statistical uncertainty in the stratified NHANES estimates and model-based HCV-related mortality estimates, we conducted a Monte Carlo simulation that respectively sampled from logit-normal and normal distributions (\( k = 10\,000 \) runs), using the standard errors for the original estimates, to produce 95% confidence intervals (CIs) for state-level estimates.

There might be state-level variability in HCV diagnosis and treatment that produces state-level variability in HCV-related deaths or proper attribution (specific ICD-10 codes) of deaths to an HCV cause, although likely limited [5]. We repeated all analyses with a broader definition of HCV-related deaths that used a combination of the HCV-specific ICD-10 codes and less-specific, more sensitive, ICD-10 codes representing cirrhosis-related and hepatocellular carcinoma–related (HCC) causes of death (Table 1) [5]. To increase specificity of these additional codes, we applied available estimates of the population attributable fraction (PAF) due to HCV infection (cirrhosis: 42%; HCC: 48%; Figure 1, equation 4) [23]. We secondarily considered other PAF estimates in less representative populations [24, 25].

We descriptively compared model findings to other publicly available reports of state estimates. Reports were excluded if HCV infection estimates solely involved applying national NHANES HCV infection prevalence estimates to total state population or if estimation was exclusively based on partial case surveillance data for HCV infection. Reports not describing the methodology used for prevalence estimation were included to facilitate more comparisons. Where possible, abstracted prevalence estimates were restricted to comparable noninstitutionalized populations.

**RESULTS**

The estimated national prevalence of anti-HCV in 2010 was 1.67% (95% CI, 1.53%–1.90%), corresponding to 3 911 800 (95% CI, 3 589 400–4 447 500) US adults with past or current HCV infection. Demographic stratum-specific estimates ranged from 0.26% for other race/ethnicity females born after 1965 to 8.02% for black males born 1945–1965 (Supplementary Table 2).

The prevalence of anti-HCV varied by state (Table 2 and Figure 2). State-specific prevalence rates ranged from 0.71% in Illinois to 3.34% in Oklahoma, with >2.5% additionally in District of Columbia, New Mexico, Oregon, and Tennessee. By census region and division, the West contained the most high-prevalence states: 10 of 13 states were above the national average, and the region-specific prevalence was 2.14%. The Pacific West division was most affected (2.30% prevalence; 865 400 persons who had been infected with HCV). The South had the second-highest anti-HCV prevalence (1.80%) and the largest number of persons with anti-HCV (n = 1 561 600). Within the South, the West South Central had the highest prevalence (2.19%), while the South Atlantic division had the most persons with anti-HCV (n = 712 900). The Midwest (1.14%) and Northeast (1.43%) had relatively lower anti-HCV prevalence. Of the 21 states in these regions, only Rhode Island had prevalence above the national average (2.12%). Although having prevalence rates lower than the national average, Michigan, New York, Pennsylvania, and Ohio each had >100 000 persons with anti-HCV.
| Region/Division/State       | 2010 Census Population | Total Persons With Anti-HCV | Anti-HCV Prevalence Rate |
|----------------------------|------------------------|----------------------------|--------------------------|
| **NORTHEAST**              |                        |                            |                          |
| New England                | 11 293 971             | 157 800                    | 1.40 (1.27–1.61)         |
| Connecticut                | 2 757 082              | 36 800                     | 1.33 (1.21–1.54)         |
| Maine                      | 1 053 828              | 11 200                     | 1.06 (0.91–1.32)         |
| Massachusetts              | 5 128 706              | 74 100                     | 1.44 (1.30–1.67)         |
| New Hampshire              | 1 029 236              | 11 000                     | 1.07 (0.91–1.33)         |
| Rhode Island               | 828 611                | 17 600                     | 2.12 (1.89–2.53)         |
| Vermont                    | 496 508                | 7200                       | 1.45 (1.23–1.89)         |
| Middle Atlantic            | 31 690 077             | 456 600                    | 1.44 (1.33–1.62)         |
| New Jersey                 | 6 726 680              | 90 700                     | 1.35 (1.23–1.53)         |
| New York                   | 15 053 173             | 223 700                    | 1.49 (1.38–1.68)         |
| Pennsylvania               | 9 910 224              | 142 100                    | 1.43 (1.30–1.62)         |
| **MIDWEST**                |                        |                            |                          |
| East North Central         | 35 269 776             | 378 800                    | 1.07 (0.98–1.22)         |
| Indiana                    | 4 875 504              | 59 100                     | 1.21 (1.08–1.40)         |
| Illinois                   | 9 701 453              | 68 400                     | 0.71 (0.64–0.80)         |
| Michigan                   | 7 539 572              | 101 200                    | 1.34 (1.22–1.52)         |
| Ohio                       | 8 805 753              | 119 000                    | 1.35 (1.22–1.53)         |
| Wisconsin                  | 4 347 494              | 31 100                     | 0.71 (0.64–0.83)         |
| West North Central         | 15 529 117             | 199 800                    | 1.29 (1.17–1.49)         |
| Iowa                       | 2 318 382              | 246 000                    | 1.06 (0.94–1.27)         |
| Kansas                     | 2 126 179              | 239 000                    | 1.41 (1.25–1.56)         |
| Minnesota                  | 4 019 862              | 41 500                     | 1.03 (0.94–1.21)         |
| Missouri                   | 4 563 491              | 76 900                     | 1.69 (1.51–1.93)         |
| Nebraska                   | 1 367 120              | 16 100                     | 1.18 (1.05–1.42)         |
| North Dakota               | 5 227 760              | 4400                       | 0.83 (0.71–1.19)         |
| South Dakota               | 6 11 383               | 6300                       | 1.03 (0.88–1.12)         |
| **SOUTH**                  |                        |                            |                          |
| South Atlantic             | 46 020 646             | 712 900                    | 1.55 (1.42–1.76)         |
| Delaware                   | 692 169                | 13 600                     | 1.97 (1.76–2.33)         |
| District of Columbia       | 500 908                | 16 400                     | 3.27 (2.87–3.90)         |
| Florida                    | 14 799 219             | 245 600                    | 1.66 (1.50–1.90)         |
| Georgia                    | 7 196 101              | 84 500                     | 1.17 (1.07–1.34)         |
| Maryland                   | 4 420 588              | 82 000                     | 1.86 (1.69–2.11)         |
| North Carolina             | 7 253 848              | 117 300                    | 1.62 (1.47–1.83)         |
| South Carolina             | 3 544 890              | 62 300                     | 1.76 (1.58–2.01)         |
| Virginia                   | 6 147 347              | 66 700                     | 1.09 (0.99–1.24)         |
| West Virginia              | 1 465 576              | 24 400                     | 1.66 (1.43–2.01)         |
| East South Central         | 14 025 119             | 264 300                    | 1.88 (1.69–2.17)         |
| Alabama                    | 3 647 277              | 52 400                     | 1.44 (1.29–1.66)         |
| Kentucky                   | 3 315 996              | 54 200                     | 1.63 (1.42–1.92)         |
| Mississippi                | 2 211 742              | 35 200                     | 1.59 (1.42–1.85)         |
| Tennessee                  | 4 850 104              | 122 500                    | 2.53 (2.25–2.91)         |
| West South Central         | 26 721 222             | 584 400                    | 2.19 (2.01–2.49)         |
| Arkansas                   | 2 204 443              | 37 500                     | 1.70 (1.50–1.99)         |
| Louisiana                  | 3 415 357              | 76 200                     | 2.23 (2.03–2.53)         |
| Oklahoma                   | 2 821 685              | 94 200                     | 3.34 (2.97–4.00)         |
| Texas                      | 18 279 737             | 378 600                    | 2.06 (1.89–2.34)         |
| **WEST**                   |                        |                            |                          |
| Mountain                   | 16 368 084             | 291 900                    | 1.78 (1.63–2.07)         |
| Arizona                    | 4 763 003              | 90 000                     | 1.89 (1.71–2.02)         |
| Colorado                   | 3 803 587              | 66 100                     | 1.74 (1.58–2.01)         |
| Idaho                      | 1 138 510              | 16 400                     | 1.44 (1.25–1.77)         |
| New Mexico                 | 1 540 507              | 42 600                     | 2.76 (2.46–3.34)         |
In sensitivity analyses that considered the addition of 2 potential indicators of HCV-related deaths, the above patterns were preserved, generally yielding state-level estimates closer to the national average (Figure 3 and Supplementary Tables 3–4). Inclusion of cirrhosis-related mortality resulted in 39 of 51 state estimates within the 95% CI bounds for the primary estimates that used only HCV-specific mortality. The median absolute deviation of state estimates between these 2 estimation approaches was 8% (interquartile range [IQR], 5%–16%, multiplicative scale). Inclusion of HCC-related mortality resulted in 50 of 51 state estimates that were within the 95% CI bounds for the primary estimates using HCV-specific mortality. The median absolute deviation of state estimates between these 2 estimation approaches was 3% (IQR, 2%–6%). Results were similar when considering alternative, but less representative PAF estimates (Supplementary Tables 5–6). There was high agreement between the primary anti-HCV prevalence estimates and 11 external estimates of state-level HCV infection prevalence (Supplementary Table 1). The 2 external estimates that utilized state-level HCV testing data (Arkansas, Oregon) were closest to our primary estimates (5.5% and 4.8% higher, respectively). To facilitate comparisons, Supplementary Tables 7–8 display model-based chronic HCV infection and NHANES standardization–based anti-HCV estimates.

DISCUSSION

We used a small-area estimation approach to synthesize data from 3 publicly available systematic, population-based data systems and provide the first comprehensive state estimates of the prevalence of anti-HCV in the United States. The estimates were robust to multiple approaches based on additional mortality data, and were comparable to independent state estimates. The method and its results can be applied by state and local health officials to guide program planning, set priorities for resource allocation, and evaluate interventions.

These state-specific estimates bring a new level of understanding to our prior knowledge about the epidemiology of HCV infection in the United States. States with high estimated anti-HCV prevalence also contain highly urban populations (eg, Rhode Island, District of Columbia) and high proportions of the population who inject drugs (eg, Vermont, Tennessee, West Virginia) or who are Native Americans (eg, Oklahoma, New Mexico; the model did not explicitly include Native American race) [26, 27]. The current anti-HCV prevalence estimates highlight states which, by their demographic structure alone, might be predicted to have lower prevalence. For example, simple NHANES standardization yields 44,000 estimated cases in Oklahoma; incorporating additional data on HCV-related deaths in Oklahoma suggests an estimated prevalence more than twice that high.

Our approach has significant methodologic strengths, as the data sources are population-based and representative of the underlying populations. However, our analysis has important limitations. We aggregated data across multiple years of NHANES to achieve a sufficient sample for stable estimates. It is possible that there were secular trends over the NHANES period. The NHANES population samples noninstitutionalized adults and therefore excludes incarcerated persons, homeless persons, those in active military service, and persons on tribal lands. These populations are critical constituents of the US HCV epidemic, comprising up to a fifth of prevalent antibody-positive persons in the United States [12]. Thus, our estimates apply only to noninstitutionalized US populations. Mortality data used to allocate the NHANES–based anti-HCV total included deaths from institutionalized populations. Complete laboratory data in NHANES were nearly universally available for the anti-HCV test, but were less complete for RNA testing.
Thus, our primary estimates of persons with anti-HCV includes the 15%–25% of people infected by HCV who may have cleared their HCV infections but retained detectable antibodies [28]. Antibody-based prevalence of individuals previously or currently infected with HCV provides a conservative overestimate of persons indicated for clinical evaluation and HCV-related medical services [8]. The HCV RNA test more accurately indicates active infection needing treatment. Importantly, the proportion of anti-HCV positive persons with detectable HCV by polymerase chain reaction is also influenced by the number of persons successfully treated and cured of their HCV infection. Over time, this proportion should decline as more persons are identified, treated, and cured; however, the proportion will vary among populations with differential access to HCV-related care and treatment.

Finally, NHANES is likely less representative of populations of persons infected in recent years. Incidence of HCV infection has been increasing since 2010, and those infected in more recent epidemics are younger, more likely to live in rural areas, and likely to acquire HCV through needle-sharing behaviors associated with opioid use [29]. Nationally, the number of recent new infections is small relative to the prevalent population, although some states with modest or low prevalence of infection have had larger increases than reflected in national
Methods that incorporate recent trends in incident HCV infection and in treatment can improve future iterations of this model. There is some mismatch between the NHANES and mortality data: mortality data represent older trends in the epidemic. Also, HCV infection may not be consistently diagnosed or recorded on death certificates by state. However, this low sensitivity of identified HCV infection in death records should not cause bias in our estimates unless the likelihood of diagnosis or recording on death certificates varies across states. We posit that deficiencies in diagnosis or recording likely occur, but are at a facility level and do not vary systematically by state. The impact of varying sensitivities of our case definition in mortality data were explored in extensive sensitivity analyses, and did not result in substantial changes to our estimates.

**CONCLUSIONS**

Although national recommendations for HCV prevention, testing, and clinical management are developed by CDC and other authorities [30], decisions regarding the capacity to deliver these services are made at the state level. State-level estimates can inform these decisions in multiple ways. First, these data can prompt reconsiderations of HCV disease burden. These modeled prevalence estimates are based on publicly available data, allowing local authorities to assess biases in data sources used in the model (ie, mortality data) and the representativeness of the modeled data in comparisons with other local data sources (ie, HCV surveillance). Second, these results, supported by local public health authorities, provide new information to engage stakeholders, resulting in agreed-upon state/local estimates of HCV infection prevalence. State-level prevalence estimates allow state and local health officials to consider more investments in surveillance and collection of other strategic data for refining our estimates. These data can also be used to revise HCV-related prevention plans and guide prevention initiatives. For example, these data can help state/local health officials estimate the number of HCV-infected persons who remain undiagnosed. The results may help state Medicaid programs to budget funds for HCV testing and treatment. More broadly, having state-level estimates calculated consistently across states will allow states to assess their standing in relation to other states and to the nation as a whole, and to adapt their
prevention and control efforts to national or other state programs that have been shown to be effective.

The approach reported here can be routinely updated as new data become available. The same approach to estimation can also be extended to yield estimates for demographic strata, areas smaller than states, and excluded populations in each state. Modeling approaches can serve as a useful method to quantify the HCV epidemic in the absence of national case surveillance data. However, this should not minimize the imperative of continuing to enhance local surveillance efforts, especially in the context of curative HCV therapies. Surveillance data may provide more accurate and reliable estimates of the burden of HCV infection and serve as the basis for public health programs to diagnose persons living with hepatitis C and link them to appropriate clinical services and treatment.

**Supplementary Data**

Supplementary materials are available at Clinical Infectious Diseases online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copylefted and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

**Notes**

**Acknowledgments.** We acknowledge Kevin Weiss for his assistance in compiling external state prevalence estimates for Supplementary Table 1 and Dr Jeremy Grey for his input on estimation methodology.

**Financial support.** We acknowledge funding from the Centers for Disease Control and Prevention (CDC) National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention Epidemic and Economic Modeling Agreement (U38 PS004646).  

**Potential conflicts of interest.** E. S. R. has received grants from CDC and personal fees from Medidata Inc and Cengage Learning. P. S. S. has received grant support through the CDC, National Institutes of Health, MAC AIDS Fund, and Gilead Sciences. T. H. S’s institution has a project funded by Gilead Sciences to develop and implement a website to help disseminate HCV surveillance and epidemiology information. All other authors report no potential conflicts. No reported conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

**References**

1. Ditah I, Ditah F, Devaki P, et al. The changing epidemiology of hepatitis C virus infection in the United States: National Health and Nutrition Examination Survey 2001 through 2010. J Hepatol 2014; 60:691–8.
2. Lingala S, Ghané MG. Natural history of hepatitis C. Gastroenterol Clin North Am 2015; 44:717–34.
3. US Department of Health and Human Services National Digestive Diseases Information Clearinghouse. Liver Transplantation. Available at: http://www.niddk.nih.gov/health-information/health-topics/liver-disease/liver-transplant/Pages/facts.aspx. Accessed 10 February 2017.
4. Ly KN, Xing J, Klevenus RM, Isles RB, Ward JW, Holmberg SD. The increasing burden of mortality from viral hepatitis in the United States between 1999 and 2007. Ann Intern Med 2012; 156:271–8.
5. Mahajan R, Xing J, Liu SJ, et al. Mortality among persons in care with hepatitis C virus infection: the Chronic Hepatitis Cohort Study (CHC-CS), 2006–2010. Clin Infect Dis 2014; 58:1055–61.
6. Spradling PR, Rupp LB, Moorman AC, et al. Chronic Hepatitis Cohort Study Investigators. Hepatitis B and C virus infection among 1.2 million persons with access to care: factors associated with testing and infection prevalence. Clin Infect Dis 2012; 55:1047–55.
7. Centers for Disease Control and Prevention. Viral Hepatitis Surveillance, United States, 2016. Available at: http://www.cdc.gov/hepatitis/statistics/2014surveillance/pdf/2014surveillance.pdf. Accessed 7 February 2017.
8. Denniston MM, Isles RB, Drobenic J, et al. Chronic hepatitis C virus infection in the United States, National Health and Nutrition Examination Survey 2003 to 2010. Ann Intern Med 2014; 160:293–300.
9. Razavi H, Elkhoury AC, Elbashir E, et al. Chronic hepatitis C virus (HCV) disease burden and cost in the United States. Hepatology 2013; 57:2164–70.
10. Centers for Disease Control and Prevention. National Notifiable Disease Surveillance System (NNDSS). Available at: http://www.cdc.gov/nndss/. Accessed 8 February 2017.
11. Zipse G, Chiappa M, Porter KS, Ostchega Y, Lewis BG, Dostal J. National Health and Nutrition Examination Survey: plan and operations, 1999–2010. Vital Health Stat 1. 2013;56:1–37.
12. Edlin BR, Eckhardt BJ, Shu MA, Holmberg SD, Swan T. Toward a more accurate estimate of the prevalence of hepatitis C in the United States. Hepatology 2015; 62:1353–63.
13. Chak E, Talal AH, Sherman KE, Schiff ER, Saab S. Hepatitis C virus infection in the United States: an estimate of true prevalence. Liver Int 2011; 31:1090–101.
14. Sokol TM, Lewis BE, Straif-Bourgeois S, Talati G, Ratard RC. Hepatitis C infection in Louisiana. J La State Med Soc 2005; 157:98–102.
15. Rao JNK, Molina I. Small area estimation. Hoboken, NJ: Wiley, 2015.
16. Lau EC, Mowat FS, Kelsh MA, et al. Use of electronic medical records (EMR) for oncology outcomes research: assessing the comparability of EMR information to patient registry and health claims data. Clin Epidemiol 2011; 3:259–72.
17. Fishman EL, Stokes A, Preston SH. The dynamics of diabetes among birth cohorts in the U.S. Diabetes Care 2014; 37:1052–9.
18. Moorman AC, Gordon SC, Rupp LB, et al. Chronic Hepatitis Cohort Study Investigators. Baseline characteristics and mortality among people in care for chronic viral hepatitis: the Chronic Hepatitis Cohort Study. Clin Infect Dis 2013; 56:40–50.
19. Centers for Disease Control and Prevention, National Center for Health Statistics. National Vital Statistics System. Available at: http://www.cdc.gov/nchs/nvss/index.htm. Accessed 7 February 2017.
20. National Association for Public Health Statistics and Information Systems. Research Requests. Available at: http://www.naphsis.org/programs/vital-statistics-data-research/ request-process. Accessed 7 February 2017.
21. US Census Bureau. Population estimates. Available at: https://www.census.gov/popest/data/datasets.html. Accessed 7 February 2017.
22. Kleinbaum DG, Kupper LL, Nizam A, Rosenberg ES. Applied regression analysis and other multivariable methods. Boston, MA: Cengage Learning, 2013.
23. Perez JF, Armstrong GL, Farrington LA, Huitin YJ, Bell BP. The contributions of hepatitis B virus and hepatitis C virus infections to cirrhosis and primary liver cancer worldwide. J Hepatol 2006; 45:529–38.
24. Makarova-Rusher OV, Altekruse SF, McNeel TS, et al. Population attributable fractions of risk factors for hepatocellular carcinoma in the United States. Cancer 2016; 122:1757–65.
25. Scaglione S, Kliethermes S, Cao G, et al. The epidemiology of cirrhosis in the United States: a population-based study. J Clin Gastroenterol 2015; 49:690–6.
26. Wise M, Bialek S, Fnedli L, Bell BR, Sorvillo F. Changing trends in hepatitis C-related mortality in the United States, 1995–2004. Hepatology 2008; 47:1128–35.
27. Van Handel MM, Rose CE, Hallisey EJ, et al. County-level vulnerability assessment for rapid dissemination of HIV or HCV infections among persons who inject drugs, United States. J Acquir Immune Defic Syndr 2016; 73:323–31.
28. Chen SL, Morgan TR. The natural history of hepatitis C virus (HCV) infection. Int J Med Sci 2006; 3:47–52.
29. Suryapradag AG, White IJ, Xu F, et al. Emerging epidemic of hepatitis C virus infections among young nonurban persons who inject drugs in the United States, 2006–2012. Clin Infect Dis 2014; 59:1411–9.
30. Centers for Disease Control and Prevention. Recommendations for prevention and control of hepatitis C virus (HCV) infection and HCV-related chronic disease. Available at: http://www.cdc.gov/hepatitis/hcv/management.htm. Accessed 11 November 2016.