2015

Toward a More Accurate Estimate of the Prevalence of Hepatitis C in the United States

Brian R. Edlin
*Weill Cornell Medical College*, bredlin.nyc@gmail.com

Benjamin J. Eckhardt
*Weill Cornell Medical College*

Marla A. Shu
*Beth Israel Medical Center*

Scott D. Holmberg
*Division of Viral Hepatitis, Centers for Disease Control and Prevention*

Tracy Swan
*Treatment Action Group, New York, NY*

Follow this and additional works at: [http://digitalcommons.unl.edu/publichealthresources](http://digitalcommons.unl.edu/publichealthresources)
Toward a More Accurate Estimate of the Prevalence of Hepatitis C in the United States

Brian R. Edlin,1,2 Benjamin J. Eckhardt,1 Marla A. Shu,3 Scott D. Holmberg,4 and Tracy Swan5

Data from the 2003-2010 National Health and Nutrition Examination Survey (NHANES) indicate that about 3.6 million people in the United States have antibodies to the hepatitis C virus, of whom 2.7 million are currently infected. NHANES, however, excludes several high-risk populations from its sampling frame, including people who are incarcerated, homeless, or hospitalized; nursing home residents; active-duty military personnel; and people living on Indian reservations. We undertook a systematic review of peer-reviewed literature and sought out unpublished presentations and data to estimate the prevalence of hepatitis C in these excluded populations and in turn improve the estimate of the number of people with hepatitis C in the United States. The available data do not support a precise result, but we estimated that 1.0 million (range 0.4 million-1.8 million) persons excluded from the NHANES sampling frame have hepatitis C virus antibody, including 500,000 incarcerated people, 220,000 homeless people, 120,000 people living on Indian reservations, and 75,000 people in hospitals. Most are men. An estimated 0.8 million (range 0.3 million-1.5 million) are currently infected. Several additional sources of underestimation, including nonresponse bias and the underrepresentation of other groups at increased risk of hepatitis C that are not excluded from the NHANES sampling frame, were not addressed in this study. Conclusion: The number of US residents who have been infected with hepatitis C is unknown but is probably at least 4.6 million (range 3.4 million-6.0 million), and of these, at least 3.5 million (range 2.5 million-4.7 million) are currently infected; additional sources of potential underestimation suggest that the true prevalence could well be higher. (Hepatology 2015;62:1353-1363)
confidence interval 3.0 million-4.2 million) had antibody to HCV, indicating past or present infection, of whom about 2.7 million (95% confidence interval 2.2 million-3.2 million) had HCV RNA-positive serum, indicating current infection.3

But while NHANES provides a wealth of valuable data on the health of the US population,1 it was designed to estimate the prevalence of conditions substantially more common than hepatitis C.4 For estimating hepatitis C prevalence, it suffers from three potential sources of underestimation. First, its sampling frame is the noninstitutionalized, housed, civilian population of the United States. By design it omits several large populations of persons at increased risk of HCV infection, including homeless persons, those in jail or prison, and those living on Indian reservations. Second, several additional groups at increased risk of hepatitis C, while not excluded from the NHANES sampling frame, are poorly represented because of small sample sizes, including Puerto Rican Americans,5 other ethnic minorities,6 and people born in high-prevalence countries.7-9 Third, nonresponse bias4 could result in underestimation if persons at elevated risk of hepatitis C differentially opt not to participate or do not provide a blood specimen.

NHANES investigators have emphasized the need to account for its omission of high-prevalence groups.3,4,10,11 To develop a more accurate estimate of the national burden of hepatitis C, we examined the first of these three potential sources of underestimation. We estimated the HCV prevalence of six populations excluded from the NHANES sampling frame (Table 1).12-19 We averaged all available estimates from the years 2003 through 2010 for each group. Hospitalized patient population estimates were obtained by multiplying the number of discharges in each available year by the average length of stay for that year and dividing by the number of days in the year; estimates for nonfederal and Veterans Affairs hospitals were added. All population size estimates reflected the number of people in each population at a single point in time, to ensure that people who move among populations are counted only once. Many more people are homeless during a given year than on a single night,13 and hepatitis C estimates that use those numbers20 are therefore severalfold higher than ours.

We then used OVID to systematically search the MEDLINE and Embase databases for articles and conference proceedings reporting the prevalence of HCV antibody in each of the six populations, limiting our search to studies conducted within 10 years of the period of the NHANES estimate (2003-2010) (Table 2). We also searched “related articles” listings, examined the references of retrieved papers, and contacted authors of published data for additional unpublished information.

Study Selection. Studies that reported HCV seroprevalence in representative or unselected samples of one of the excluded population groups in the United States were included; those which sampled subjects at selectively higher risk by virtue of risk factors, symptoms, or requests for testing were excluded.

Data Extraction and Quality Assessment. One investigator extracted the data from each study, and at least one additional investigator verified the accuracy of the data. The dates of testing, number of persons tested, number of persons testing HCV antibody-positive, and HCV seroprevalence were extracted. Authors were contacted to obtain data missing from published reports.

### Materials and Methods

Data Sources and Searches. We used publicly available estimates of the size of each of the six population groups excluded from the NHANES sampling frame (Table 1).12-19 We averaged all available estimates from the years 2003 through 2010 for each group. Hospitalized patient population estimates were obtained by multiplying the number of discharges in each available year by the average length of stay for that year and dividing by the number of days in the year; estimates for nonfederal and Veterans Affairs hospitals were added. All population size estimates reflected the number of people in each population at a single point in time, to ensure that people who move among populations are counted only once. Many more people are homeless during a given year than on a single night,13 and hepatitis C estimates that use those numbers20 are therefore severalfold higher than ours.

### Study Selection

Studies that reported HCV seroprevalence in representative or unselected samples of one of the excluded population groups in the United States were included; those which sampled subjects at selectively higher risk by virtue of risk factors, symptoms, or requests for testing were excluded.

Data Extraction and Quality Assessment. One investigator extracted the data from each study, and at least one additional investigator verified the accuracy of the data. The dates of testing, number of persons tested, number of persons testing HCV antibody-positive, and HCV seroprevalence were extracted. Authors were contacted to obtain data missing from published reports.

### Table 1. Estimated Size of Population Groups Excluded from NHANES

<table>
<thead>
<tr>
<th>Population</th>
<th>Estimated Size</th>
<th>Source</th>
<th>Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nursing home residents</td>
<td>1,446,999</td>
<td>National Nursing Home Survey16</td>
<td>2004</td>
</tr>
<tr>
<td>Active-duty military</td>
<td>1,404,060</td>
<td>US Census Bureau17</td>
<td>2009</td>
</tr>
</tbody>
</table>

Abbreviations: CDC/NCHS, Centers for Disease Control and Prevention/National Center for Health Statistics.
For the incarcerated population, available seroprevalence estimates for each state were averaged and a national estimate was calculated by weighting the state averages by the size of each state’s incarcerated population. Where separate estimates were available for men and women, they were weighted according to the sex ratio of prisoners for the respective state. For each of the other groups, we calculated the mean seroprevalence of the available studies, weighting each study by its sample size. We calculated binomial confidence intervals for the seroprevalence estimate from each study where available data permitted. To conservatively account for interstudy variability within each of the six population groups, we used the range of the point estimates from the separate studies to represent the uncertainty in our estimate. The size of each population group was multiplied by its estimated seroprevalence, and the totals were summed and added to the 2003-2010 NHANES estimate for the number of US residents with HCV antibody. To estimate the number of persons with current HCV infection, we used the NHANES finding that 82% of antibody-positive persons have HCV RNA. The same proportion, 82%, was observed in a high-prevalence sample of people who inject drugs. This analysis updates an earlier one that used the same methods.

Data Synthesis and Analysis. For the incarcerated population, available seroprevalence estimates for each state were averaged and a national estimate was calculated by weighting the state averages by the size of each state’s incarcerated population. Where separate estimates were available for men and women, they were weighted according to the sex ratio of prisoners for the respective state. For each of the other groups, we calculated the mean seroprevalence of the available studies, weighting each study by its sample size. We calculated binomial confidence intervals for the seroprevalence estimate from each study where available data permitted. To conservatively account for interstudy variability within each of the six population groups, we used the range of the point estimates from the separate studies to represent the uncertainty in our estimate. The size of each population group was multiplied by its estimated seroprevalence, and the totals were summed and added to the 2003-2010 NHANES estimate for the number of US residents with HCV antibody. To estimate the number of persons with current HCV infection, we used the NHANES finding that 82% of antibody-positive persons have HCV RNA. The same proportion, 82%, was observed in a high-prevalence sample of people who inject drugs. This analysis updates an earlier one that used the same methods.

Results

The sizes of the six population groups ranged from 478,054 to 2,186,230 (Table 1). The search for seroprevalence data yielded 2828 unique articles, of which 36 met criteria for inclusion (Table 2), with seven additional studies published in abstract form only. One study of prisoners and two studies of hospitalized patients were excluded because they were conducted >10 years before 2003.

Sixteen published studies and six studies published in abstract form only reported HCV seroprevalence among persons in penal institutions in 23 states (Table 3). The studies differed in sample design and methodology. Most, but not all, sampled persons as they entered these institutions. The weighted mean prevalence of HCV antibody among all persons in all studies combined was 23.1%. The prevalences in the 22 studies ranged from 7.5% through 44.0%.

Nine studies of HCV seroprevalence among homeless persons have been published (Table 4). The study design and selection criteria differed in each study. The weighted mean prevalence of HCV antibody among all persons in all studies combined was 32.1%. The prevalences in the nine studies ranged from 7.5% through 52.5%.

Seven studies of HCV seroprevalence among persons in hospitals have been published, and an additional study was published in abstract form only (Table 4). One studied source patients of needle-stick injuries; two were conducted in emergency departments and two in a psychiatric hospital, with HCV antibody prevalences between 4.0% and 38% in all the studies. The weighted mean prevalence of HCV antibody among all persons in these studies combined was 15.6%.

One small study in nursing home residents demonstrated an HCV seroprevalence of 4.5%. A random sample of 10,000 active-duty military personnel found an HCV antibody prevalence of 0.48%, and a smaller study of military blood donors showed a seroprevalence of 0.84%. A single study of Native Americans served by the Indian Health Service reported an HCV seroprevalence of 11.5% (Table 4).

When we multiplied the estimated seroprevalence for each of the six populations by the population size, we projected that in 2003-2010 505,350 incarcerated people, 222,100 homeless individuals, 74,576 hospitalized patients, 65,113 nursing home residents, 7020 active military personnel, and 123,224 persons living on Indian reservations had HCV antibody. Summing these, we estimated that 997,384 (range 355,466-1,813,661) persons in the United States had HCV antibody in addition to the NHANES estimate (Table 5). Of these, an estimated 817,855 (range 291,482-1,487,202) are currently infected.

<table>
<thead>
<tr>
<th>Population</th>
<th>Search Terms</th>
<th>Medline Articles</th>
<th>Embase Articles</th>
<th>Unique Articles</th>
<th>Full Text Articles</th>
<th>Included Articles</th>
<th>Included Articles</th>
<th>Included Articles</th>
<th>Included Articles</th>
<th>Included Articles</th>
<th>Included Articles</th>
<th>Included Articles</th>
<th>Included Articles</th>
<th>Included Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incarcerated</td>
<td>&quot;hepatitis C&quot; and &quot;prison&quot; or &quot;jail&quot; or &quot;incarceration&quot;</td>
<td>651</td>
<td>766</td>
<td>879</td>
<td>51</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homeless</td>
<td>&quot;hepatitis C&quot; and &quot;homeless&quot;</td>
<td>173</td>
<td>223</td>
<td>245</td>
<td>31</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hospitalized</td>
<td>&quot;hepatitis C&quot; and &quot;hospitalized&quot; or &quot;emergency room&quot; or &quot;emergency ward&quot; or &quot;inpatient&quot;</td>
<td>559</td>
<td>1252</td>
<td>1340</td>
<td>31</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nursing homes</td>
<td>&quot;hepatitis C&quot; and &quot;nursing home,&quot; &quot;nursing facility,&quot; or &quot;long-term care facility&quot; or &quot;skilled nursing facility&quot;</td>
<td>13</td>
<td>38</td>
<td>38</td>
<td>12</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Military</td>
<td>&quot;hepatitis C&quot; and &quot;military,&quot; &quot;navy,&quot; &quot;air force,&quot; &quot;army,&quot; or &quot;marines&quot;</td>
<td>172</td>
<td>207</td>
<td>228</td>
<td>15</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Native Americans</td>
<td>&quot;hepatitis C&quot; and &quot;Native American,&quot; &quot;Indian, North American&quot; or &quot;Indian reservation&quot;</td>
<td>55</td>
<td>74</td>
<td>98</td>
<td>19</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>1623</td>
<td>2560</td>
<td>2828</td>
<td>159</td>
<td>36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Adding these numbers to the most recent NHANES estimate suggests that at least 4.6 million people in the United States (range 3.4 million-6.0 million) have HCV antibody (Table 6). Of these, we project that at least 3.5 million (range 2.5 million-4.7 million) are currently infected (Table 6).

Sex-specific prevalence data allowed us to estimate the proportions of men and women among the infected in three of the excluded populations (see Appendix A). Men predominated in the excluded populations overall, as well as among the infected persons in those populations. Men comprised 87.8% of infected prisoners, 74.5% of infected homeless persons, and 90.9% of infected military personnel compared with 64% of infected persons in the NHANES population. Thus, NHANES appears to underestimate the numbers of HCV-infected men more than women.

**Discussion**

Estimates of disease prevalence are important for establishing disease burdens, identifying health disparities, guiding policy, targeting interventions, and allocating resources. This analysis highlights several challenges to estimating the national prevalence of hepatitis C. First, while household surveys contribute invaluable information about population health, to be accurate their findings must be adjusted to account for populations not sampled or not well represented in the sample. NHANES uses consistent, standardized methodology that allows monitoring of trends over time but yields prevalence estimates that underestimate the total burden of hepatitis C in the United States. Figures derived from NHANES appear to underestimate the US HCV seroprevalence by at least 1 million persons. Probably at least 4.6 million Americans have HCV antibody and at least 3.5 million are currently infected.
Second, household surveys underestimate the prevalence of conditions concentrated in disenfranchised populations and, therefore, underestimate the extent of health disparities. The largest contributions to the hepatitis C underestimate are from prisoners, homeless persons, and residents of Indian reservations—marginalized groups facing social stigma, economic disadvantage, elevated rates of comorbidities, severe health disparities, and reduced access to high-quality health care. These groups also include a disproportionately high representation of ethnic minorities affected by health disparities and ethnic bias. It is particularly important that these vulnerable populations are not overlooked when allocating resources, designing interventions, and planning for health care needs.

Third, this analysis reveals the paucity of reliable data on the prevalence of hepatitis C in populations known to be at increased risk of the disease. A number of population groups have been excluded from the National Health and Nutrition Examination Survey (NHANES), which is the primary source of prevalence data in the United States. These groups are estimated to include a significant number of individuals with hepatitis C.

### Table 4. Hepatitis C Seroprevalence Studies in Homeless, Hospitalized, Nursing Home, Military, and Native American Populations

<table>
<thead>
<tr>
<th>Reference</th>
<th>Location</th>
<th>Study Dates</th>
<th>Total No. Tested</th>
<th>No. HCV Antibody-Positive</th>
<th>HCV Antibody Prevalence</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Desai et al.</td>
<td>Bedford, MA</td>
<td>1993-1998</td>
<td>418</td>
<td>184</td>
<td>44%</td>
<td>39.2%-48.9%</td>
</tr>
<tr>
<td>Nymathi et al.</td>
<td>Los Angeles, CA</td>
<td>1995-1999</td>
<td>884</td>
<td>197</td>
<td>22%</td>
<td>19.6%-25.2%</td>
</tr>
<tr>
<td>Cheung et al.</td>
<td>Palo Alto, CA</td>
<td>1995-2000</td>
<td>787</td>
<td>314</td>
<td>39.9%</td>
<td>36.5%-43.4%</td>
</tr>
<tr>
<td>Rosenblum et al.</td>
<td>New York, NY</td>
<td>1997-1998</td>
<td>139</td>
<td>45</td>
<td>32%</td>
<td>24.7%-40.8%</td>
</tr>
<tr>
<td>Schwarz et al.</td>
<td>Baltimore, MD</td>
<td>2001-2004</td>
<td>168</td>
<td>32</td>
<td>19%</td>
<td>13.4%-25.8%</td>
</tr>
<tr>
<td>Stein and Nymathi</td>
<td>Los Angeles, CA</td>
<td>2002-2003</td>
<td>198</td>
<td>104</td>
<td>52.5%</td>
<td>45.3%-59.6%</td>
</tr>
<tr>
<td>Gelberg et al.</td>
<td>Los Angeles, CA</td>
<td>2003-2004</td>
<td>534</td>
<td>22%</td>
<td>19.8%</td>
<td>15.9%-24.7%</td>
</tr>
<tr>
<td>Strehlow et al.</td>
<td>8 cities</td>
<td>2003-2004</td>
<td>387</td>
<td>120</td>
<td>31.0%</td>
<td>26.4%-35.9%</td>
</tr>
<tr>
<td>Boyce et al.</td>
<td>Hawaii</td>
<td>2006</td>
<td>40</td>
<td>3</td>
<td>7.5%</td>
<td>1.6%-20.4%</td>
</tr>
</tbody>
</table>

**Weighted mean prevalence:** 32.1% Range: 7.5%-52.5%

### Table 5. Estimated Numbers of HCV Infections in Population Groups Excluded From NHANES

<table>
<thead>
<tr>
<th>Population</th>
<th>Estimated Size</th>
<th>HCV Prevalence</th>
<th>Number of Infected Persons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incarcerated</td>
<td>2,186,230</td>
<td>21.3%</td>
<td>505,350</td>
</tr>
<tr>
<td>Homeless</td>
<td>691,899</td>
<td>32.1%</td>
<td>222,100</td>
</tr>
<tr>
<td>Hospitalized</td>
<td>478,054</td>
<td>15.6%</td>
<td>74,576</td>
</tr>
<tr>
<td>Nursing homes</td>
<td>1,446,959</td>
<td>4.5%</td>
<td>65,113</td>
</tr>
<tr>
<td>Military</td>
<td>1,404,060</td>
<td>0.5%</td>
<td>7020</td>
</tr>
<tr>
<td>Indian reservations</td>
<td>1,069,411</td>
<td>11.5%</td>
<td>123,224</td>
</tr>
<tr>
<td>Total</td>
<td>997,384</td>
<td></td>
<td>355,466-1,813,661</td>
</tr>
</tbody>
</table>
limitations of the available data impair the accuracy of the estimates presented here. First, there are no representative samples of five of the six populations examined. The seroprevalence studies we used, with the exception of a single study of active-duty military personnel, were not designed to recruit nationally representative samples. Second, the size of two of the populations is not known with accuracy. There is no precise estimate of the number of homeless persons in the United States, and the US census faces substantial challenges in enumerating the Native American population on reservations. Third, HCV seroprevalence studies were available from jails and prisons in only 23 states and no federal prisons. While these states cannot be considered representative of the United States as a whole, it is worth noting that taken together they account for 55% of the nation’s prisoners. Fourth, HCV seroprevalence estimates were available for only nine homeless samples and eight hospitals, including several Veterans Affairs and psychiatric hospitals, which may have higher rates of HCV seroprevalence than other hospitals. Fifth, only a single clinic-based study of Native Americans was available, which might have overestimated the hepatitis C prevalence in that population. Surveillance data suggest, however, that the prevalence of hepatitis C on Indian reservations is elevated. It should also be noted that we did not examine the ethnic composition of the undercounted persons we report and that our methods cannot be used to monitor trends over time. Finally, studies showing higher prevalence rates may have been more likely to be published, which would have biased our estimates upward.

But balanced against these possible sources of overestimation are several reasons the current study may still underestimate the number of infected persons. First, persons in temporary or unstable housing outnumber those who are literally homeless on any given night by severalfold. For example, an estimated 6.8 million people were “doubled-up”—i.e., living with others—in 2010. Persons not in their “usual place of residence” were not included in NHANES. While the HCV seroprevalence of this precariously housed population is not known, these persons could conservatively account for another half-million to 1 million or more additional persons with HCV antibody.

Second, nearly all of the prison studies cited in our study sampled new entrants to prison, who are on average younger than the overall prison population and include fewer members of the higher-prevalence birth cohorts. HCV seroprevalence in the United States peaked among persons born during 1945-1965. Adjusting seroprevalence of prisoners for this difference, an average of about 3 years, would add 2.45% to the seroprevalence estimate, or about 50,000 additional infected persons (see Appendix B).

Third, nearly one-third of persons sampled by NHANES were not interviewed or did not provide a blood specimen. The NHANES HCV seroprevalence estimate is based on only 381 positive antibody tests, while 13,824 sampled persons were not interviewed or did not provide a blood sample. If even 1%-2% of these unsampled persons had injected illicit drugs unsafely and did not want to be asked questions about their health behavior, reveal needle tracks, undergo what may have been a painful or embarrassing phlebotomy experience, or participate in government research at all, the true prevalence could be one-third to two-thirds higher than the NHANES estimate. In addition, NHANES phlebotomists reported that at times they were unable to obtain a blood sample because prior illicit injection drug use made venipuncture too difficult. A study of nonresponse bias in a household survey of human immunodeficiency virus (HIV) seroprevalence illustrates the potential magnitude of this effect. Persons who declined to participate in the seroprevalence survey (but consented to be interviewed in a subsequent study) were more than twice as likely as survey responders to report a history of injection drug use. Of note, adjustment for nonresponse of this magnitude still resulted in an underestimate of the HIV seroprevalence compared with estimates determined by other methods; backcalculation models yielded estimates 1.8-fold to 4.6-fold higher than the estimate derived from the household survey even after nonresponse bias was accounted for.

Fourth, while NHANES oversamples African Americans and Mexican Americans to improve representation of

<table>
<thead>
<tr>
<th>Estimate</th>
<th>HCV Antibody-Positive Number (Range)</th>
<th>HCV RNA-Positive Number (Range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NHANES (2003-2010)</td>
<td>3.6 million (3.0 million-4.2 million)</td>
<td>2.7 million (2.2 million-3.2 million)</td>
</tr>
<tr>
<td>Added populations</td>
<td>1.0 million (0.4 million-1.8 million)</td>
<td>0.8 million (0.3 million-1.5 million)</td>
</tr>
<tr>
<td>Total</td>
<td>4.6 million (3.4 million-6.0 million)</td>
<td>3.5 million (2.5 million-4.7 million)</td>
</tr>
</tbody>
</table>
these populations, it is unable to adequately sample other groups who may have elevated hepatitis C prevalence, including other ethnic minority groups and people born in high-prevalence countries. \(^5\) \(^9\) Thus, although these groups are not excluded from the NHANES sampling frame, they are nonetheless poorly represented in the survey. \(^5\) Hepatitis C prevalence is elevated among Puerto Rican Americans \(^5\) and immigrants from Egypt \(^7\) and Somalia, \(^8\) for example; and in one study of immigrants from the former Soviet Union, of whom some 2.9 million are living in the United States, the HCV seroprevalence was 28.3%. \(^7\) The seroprevalence among immigrants from other high-prevalence countries \(^77\) is unknown.

Fifth, all household surveys miss people, even among groups that are not poorly represented or excluded from the sampling frame. Studies of the 2000 US census coverage, for example, estimated that \(\sim 3\) million to 4 million people were uncounted. \(^78\) \(^79\) Uncounted persons are disproportionately young, single, male, poor, of minority ethnicity, and therefore on average more likely to have hepatitis C. The coverage of national household surveys such as NHANES, with only a fraction of the resources available to the US census to reach disadvantaged groups, is likely to be at least as impaired in this regard.

Finally, NHANES has limited ability to discern new transmission patterns concentrated in specific foci in population subgroups. Increased HCV transmission is occurring in at least 30 states, especially in rural and suburban areas, on the heels of dramatic increases in opioid use among young adults during the past decade. \(^80\) \(^83\) These new infections have yet to register in NHANES.

For these reasons, the findings presented here may still underestimate the true number of HCV-infected persons in the United States. The present study examined only one of several sources of potential underestimation. Further work is needed to assess the contributions of these other possible sources of error in the estimate, as has been done for HIV. \(^84\) An analysis of the NHANES HIV prevalence estimate concluded that the true HIV seroprevalence was 1.4-fold to 2.0-fold higher than the NHANES estimate because of nonresponse bias and the exclusion of high-risk populations. \(^4\) \(^85\) If HCV seroprevalence was underestimated by a similar proportion—and there is little reason to think it was not—the true seroprevalence of HCV would be 5 million to 7 million.

This study illustrates the limitations of household surveys for ascertaining sequelae of stigmatized behavior. Injection drug use is highly prevalent in prisons and marginally housed populations because our nation criminalizes and incarcerates people who inject drugs and disqualifies them from receiving public aid such as housing assistance. And because we stigmatize illicit drug use, those who have used drugs and are free and housed are understandably reluctant to disclose it. Thus, the National Household Survey on Drug Abuse estimated that 440,000 persons in the United States had injected illicit drugs in the past year, while contemporaneous estimates from other sources put the figure at 1.5 million to 2.0 million. \(^86\) \(^87\) Over 1 million people who injected drugs were either missed or misclassified in that household survey. Because 50%-90% of people who currently inject illicit drugs have HCV antibody (depending on the duration of their use), \(^88\) \(^89\) this provides further support for the likelihood that among current injectors alone, a million more Americans have been infected with HCV than might be detected by a household survey.

These limitations underscore the need for better assessment and monitoring of the health needs of socially marginalized populations. This problem impairs public health responses to nearly all diseases, not just hepatitis C; surveillance systems are not accurate if they overlook disenfranchised groups. \(^5\) Hepatitis C prevalence estimates should not be based on NHANES alone but should be augmented with expanded HCV screening, case surveillance, and focused seroprevalence studies of established and emerging groups at elevated risk. These testing initiatives can provide data for surveillance purposes at the same time as they identify infected persons and link them to care and treatment.

For the meantime, projections based on NHANES data may underestimate current health disparities and the burden of liver disease that can be expected in the coming decades. Most models projecting the future disease burden use NHANES estimates without correction for the excluded populations or other sources of underestimation. \(^90\) \(^93\) As new highly effective antiviral regimens offer the hope of transforming the hepatitis C epidemic, it will be important to plan appropriately. \(^89\) \(^94\)

These findings may heighten concern about the affordability of providing antiviral treatment to all who need it. Only an estimated 5%-6% of infected people in the United States have been successfully treated. \(^11\) It is important to realize, however, that curative antiviral treatment, and the associated costs, will occur over many years. Most people with hepatitis C do not yet know their status, \(^11\) \(^95\) and many of the groups discussed in this article have substantial barriers to health care access. \(^94\) Resources for addressing these deficiencies have not kept pace with the need. \(^39\) It will take years of concerted effort before most infected people, especially...
those in underserved groups, are tested, engaged in care, and referred for treatment. Indeed, a national commitment will be needed to bring screening, treatment, and prevention services to these populations.\textsuperscript{11,89,94,95}

Currently, however, public and private payers are limiting access to the new, high-priced oral drugs for hepatitis C, requiring abstinence from drugs and alcohol and the presence of advanced liver disease,\textsuperscript{96,97} which means that HCV will continue to spread among the highly affected groups most in need, such as young people who inject drugs.\textsuperscript{80,81} More accurate estimates of hepatitis C prevalence and incidence can help inform the planning and negotiation of strategies to reach, test, treat, cure, and prevent every case of hepatitis C.\textsuperscript{94}

**Appendix A: Sex Ratio Methods**

Where sex-specific prevalence rates were available, we calculated the M:F ratio of infected persons by multiplying the M:F prevalence in each study with sex-specific prevalence data\textsuperscript{20,28,32-36,38,41-45,49,52,53,64,65} by the sex ratio of the population:

\[
\text{M:F ratio of infected persons} = \frac{\text{No. infected women in population}}{\text{No. infected men in population}} = \frac{\text{HCV prevalence in women}}{\text{HCV prevalence in men}} \times \frac{\text{Total no. women in population}}{\text{Total no. men in population}}
\]

For prison studies we used the sex ratio of the respective state prison system. For homeless studies we used the estimated overall national sex ratio.

We then converted the M:F ratio of infected persons to the percentage of men over the percentage of women. For prison and homeless studies we averaged the percentages, weighting prison studies by the total prison population for the respective state and homeless studies by the sample size of the study. For the military population, we used the one representative national study\textsuperscript{64} (self-weighted with respect to sex ratios) and calculated the M:F ratio of infected persons directly.

**Appendix B: Projected Seroprevalence of People Living in Prison Versus People Entering Prison**

**Background.** Entrants to prison are younger than the overall prison population. This is true not only because persons age while they serve time in prison but also because prison populations are enriched with people serving longer terms, in proportion to the length of time they serve. The mean age of people entering prison in the United States in 2012 was 33 years, while the mean age of people living in prison was 36 years.\textsuperscript{98} Thus, because HCV prevalence is strongly associated with older age in people <65 years (the age of 98% of prisoners), the HCV prevalence of prison entrants underrepresents the HCV prevalence of people serving time in prison.

**Methods and Results.** To determine the effect on HCV seroprevalence of the age difference between people entering prison and people living in prison, we used published data on HCV prevalence rates by birth cohort of entrants to prison in Pennsylvania.\textsuperscript{31} The age distribution of entrants to prison in Pennsylvania was similar to that of entrants to prison nationwide.\textsuperscript{98} While the HCV prevalence of persons entering prison in Pennsylvania might not represent the prevalence of persons entering prison nationwide, we used the distribution of rates by birth cohort as a proxy for the distribution by birth cohort nationwide. (For example, the peak prevalences in Pennsylvania entrants were in the 1945-1965 birth cohort, similar to the national household survey data).\textsuperscript{10} When we applied the birth cohort–specific prevalence data from Pennsylvania to the age distribution of entrants in 2012 nationwide, we obtained an overall HCV prevalence similar to that reported from Pennsylvania (17.9% versus 18.1%). When we then applied the same prevalence data to the age distribution of persons serving time in prison in 2012 nationwide, the calculated prevalence was 20.3%. The difference in projected HCV prevalence between the entrants (17.85%) and persons serving time (20.30%) was 2.45%.

**Summary.** To adjust the projected seroprevalence of prisoners for the difference between the age of people entering prison and the age of people living in prison, we applied the birth cohort–specific HCV prevalences of persons entering prison in Pennsylvania\textsuperscript{11} to the age distributions of persons entering prison and persons living in prison nationwide. The projected prevalence of persons serving time was 2.45% greater than the projected prevalence among entrants.

**References**


40. Smith L, Wang L, Wright L, Sabin K, Gilebatis D, Smith P. Hepatitis C virus (HCV) seroprevalence in incoming inmates in New York State
56. Pham D, Walshe D, Montgomery J, Buskell-Bales Z, Collier K. Seroe-

55. Austin GE, Jensen B, Leete J, De l'Aune W, Bhatnagar J, Racine M,

54. Boyce DE, Tice AD, Ona F, Akinaka K, Lusk H. Viral hepatitis in a


52. Gelberg L, Robertson MJ, Arangua L, Leake BD, Sumner G, Moe A,

51. Stein JA, Nyamathi A. Correlates of hepatitis C virus infection in

50. Schwarz KB, Garrett B, Alter MJ, Thompson D, Strathdee SA. Sero-

47. Nyamathi AM, Dixon EL, Wiley D, Christiani A, Lowe A. Hepatitis

46. Desai RA, Rosenheck RA, Agnello V. Prevalence of hepatitis C virus

43. Macalino GE, Vlahov D, Sanford-Colby S, Patel S, Sabin K, Salas C,

41. Larney S, Mahowald MK, Scharff N, Flanigan TP, Beckwith CG,

40. Hall MR, Ray D, Payne J. Prevalence of hepatitis C, hepatitis B, and

38. Schwarz KB, Garrett B, Pierre JM, Guzik LH, Kisicki MD, Wirshing D,

37. Brillman JC, Crandall CS, Florence CS, Jacobs JL. Prevalence and risk

36. Zaller ND. Epidemiology of hepatitis C virus in Pennsylvania state


34. Montana 2008 annual report: STDs including HIV/AIDS and hepatitis

33. Burgard S, Seefeldt KS, Zelner S. Housing instability and health: find-

32. Division of Viral Hepatitis. Viral Hepatitis Surveillance, United States,

31. Stein JA, Nyamathi A. Correlates of hepatitis C virus infection in homeless

30. Schwarz KB, Garrett B, Pierre JM, Guzik LH, Kisicki MD, Wirshing D,


28. Cheung RC, Hanson AK, Maganti K, Keeffe EB, Matsui SM. Viral

27. Calore BL, Cheung RC, Giori NJ. Prevalence of hepatitis C virus infection in the veteran population undergoing total joint arthroplasty. 

Bisceglie AM. Seroprevalence of viral hepatitis in an older nursing


24. Baillargeon J, Wu H, Kelley MJ, Grady J, Linthicum L, Dunn K. Hepa-

23. Macalino GE, Vlahov D, Sanford-Colby S, Patel S, Sabin K, Salas C, 


"human immunodeficiency virus in a Grand Rapids, Michigan emerg-

20. Meyer JM. Prevalence of hepatitis A, hepatitis B, and HIV among hep-

19. Pham D, Walshe D, Montgomery J, Buskell-Bales Z, Collier K. Sero-
pidemiology of hepatitis C and B in an urban VA medical center. 
HEPATOLOGY 1994;20(Suppl.):236A.

18. Macalino GE, Vlahov D, Sanford-Colby S, Patel S, Sabin K, Salas C, 

17. Larney S, Mahowald MK, Scharff N, Flanigan TP, Beckwith CG,

16. Macalino GE, Vlahov D, Sanford-Colby S, Patel S, Sabin K, Salas C, 

15. Larney S, Mahowald MK, Scharff N, Flanigan TP, Beckwith CG,

14. Larney S, Mahowald MK, Scharff N, Flanigan TP, Beckwith CG,

13. Larney S, Mahowald MK, Scharff N, Flanigan TP, Beckwith CG,

12. Larney S, Mahowald MK, Scharff N, Flanigan TP, Beckwith CG,

11. Larney S, Mahowald MK, Scharff N, Flanigan TP, Beckwith CG,

10. Larney S, Mahowald MK, Scharff N, Flanigan TP, Beckwith CG,

9. Larney S, Mahowald MK, Scharff N, Flanigan TP, Beckwith CG,

8. Larney S, Mahowald MK, Scharff N, Flanigan TP, Beckwith CG,


