RESEARCH ARTICLE

Estimating HIV, HCV and HSV2 incidence from emergency department serosurvey [version 1; peer review: awaiting peer review]

Simon E.F. Spencer ^{1*}, Oliver Laeyendecker^{2,3*}, Louise Dyson ^{4,5}, Yu-Hsiang Hsieh³, Eshan U. Patel³, Richard E. Rothman³, Gabor D. Kelen³, Thomas C. Quinn^{2,3}, T. Deirdre Hollingsworth ⁶

¹Department of Statistics, University of Warwick, Coventry, CV7 4AL, UK

²Laboratory of Immunoregulation, Division of Intramural Research, NIAID, NIH, Baltimore, MD, USA

³Johns Hopkins University, Baltimore, MD, USA

⁴Mathematics Institute, University of Warwick, Coventry, CV4 7AL, UK

⁵School of Life Sciences, University of Warwick, Coventry, UK

⁶Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford, UK

* Equal contributors

V1 First published: 02 Aug 2021, 5:116
<https://doi.org/10.12688/gatesopenres.13261.1>

Latest published: 02 Aug 2021, 5:116
<https://doi.org/10.12688/gatesopenres.13261.1>

Open Peer Review

Approval Status Awaiting Peer Review

Any reports and responses or comments on the article can be found at the end of the article.

Abstract

Background: Our understanding of pathogens and disease transmission has improved dramatically over the past 100 years, but coinfection, how different pathogens interact with each other, remains a challenge. Cross-sectional serological studies including multiple pathogens offer a crucial insight into this problem.

Methods: We use data from three cross-sectional serological surveys (in 2003, 2007 and 2013) in a Baltimore emergency department to predict the prevalence for HIV, hepatitis C virus (HCV) and herpes simplex virus, type 2 (HSV2), in a fourth survey (in 2016). We develop a mathematical model to make this prediction and to estimate the incidence of infection and coinfection in each age and ethnic group in each year.

Results: Overall we find a much stronger age cohort effect than a time effect, so that, while incidence at a given age may decrease over time, individuals born at similar times experience a more constant force of infection over time.

Conclusions: These results emphasise the importance of age-cohort counselling and early intervention while people are young. Our approach adds value to data such as these by providing age- and time-specific incidence estimates which could not be obtained any other way, and allows forecasting to enable future public health planning.

Keywords

Epidemiology, coinfection, serology, Bayesian statistics, repeated cross-sectional studies

Corresponding authors: Simon E.F. Spencer (s.e.f.spencer@warwick.ac.uk), Oliver Laeyendecker (olaeyen1@jhmi.edu)

Author roles: **Spencer SEF:** Conceptualization, Formal Analysis, Investigation, Methodology, Software, Writing – Original Draft Preparation, Writing – Review & Editing; **Laeyendecker O:** Conceptualization, Data Curation, Formal Analysis, Investigation, Writing – Original Draft Preparation, Writing – Review & Editing; **Dyson L:** Conceptualization, Formal Analysis, Investigation, Methodology, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Hsieh YH:** Data Curation, Investigation, Writing – Review & Editing; **Patel EU:** Data Curation, Investigation, Writing – Review & Editing; **Rothman RE:** Data Curation, Funding Acquisition, Investigation, Writing – Review & Editing; **Kelen GD:** Data Curation, Investigation, Writing – Review & Editing; **Quinn TC:** Data Curation, Funding Acquisition, Investigation, Writing – Review & Editing; **Hollingsworth TD:** Conceptualization, Investigation, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was supported in part by the Division of Intramural Research, National Institute of Allergy and Infectious Diseases (NIAID). Other support was provided by NIH grants (K01AI100681, R01-095068). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH. SEFS gratefully acknowledges funding by MRC grant MR/P026400/1 and EPSRC grant EP/R018561/1. TDH and SEFS were supported by the Bill and Melinda Gates Foundation through the NTD Modelling Consortium (OPP1184344).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2021 Spencer SEF *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The author(s) is/are employees of the US Government and therefore domestic copyright protection in USA does not apply to this work. The work may be protected under the copyright laws of other jurisdictions when used in those jurisdictions.

How to cite this article: Spencer SEF, Laeyendecker O, Dyson L *et al.* **Estimating HIV, HCV and HSV2 incidence from emergency department serosurvey [version 1; peer review: awaiting peer review]** Gates Open Research 2021, 5:116 <https://doi.org/10.12688/gatesopenres.13261.1>

First published: 02 Aug 2021, 5:116 <https://doi.org/10.12688/gatesopenres.13261.1>

Introduction

Currently there are an estimated 38 million individuals infected with HIV¹, while worldwide an estimated 71 million individuals are seropositive for HCV². Estimates for HSV2 are even higher, at 400 million individuals³. Despite these high prevalences, estimating HIV, HCV and HSV2 incidence using current methods is challenging, and this difficulty is exacerbated by the initially asymptomatic nature of most of these infections. Incidence estimates are critical because they determine the current position of the leading edge of the epidemic and form part of WHO elimination targets^{4,5}. In addition, at the population level, an accurate estimate of disease incidence allows countries to determine their future healthcare needs and assess the impact of prevention efforts.

There are a number of ways to estimate the population-level incidence of these diseases. The gold standard for incidence estimation is longitudinally-followed cohorts, measuring the rate of seroconversion by follow-up time. However, such longitudinal cohorts are expensive to maintain and suffer from selection bias and the Hawthorne effect⁶. Biomarker-based approaches also exist for measuring the incidence of HIV^{7,8}, HCV^{9,10}, and HSV2¹¹ via cross-sectional surveys. Although cross-sectional, biomarker-based methods have been applied to estimate HIV incidence in a number of settings^{12–14}, the HCV and HSV2 biomarkers have not been validated as tools for estimating incidence at the population level. Therefore, cross-sectional studies of prevalence are still a standard tool for routine surveillance of these diseases^{15,16}. These studies are still susceptible to a host of problems, including bias in the people surveyed, differential survival rates, and lack of information on anti-retroviral treatment status, but they are relatively simple to perform and provide an important insight despite studying a necessarily anonymous dataset.

Historically, the Johns Hopkins Hospital Emergency Department (JHH ED) has conducted serial identity-unlinked serosurveys to monitor the HIV epidemic among the marginalized inner-city populations of Baltimore, Maryland. These surveys demonstrated a high burden of HSV2, HIV and HCV, particularly among African Americans^{17–20}. Previously they have been used to determine the care continuum among HIV infected individuals²¹, and also used to evaluate the recommended HCV testing guidelines²². These datasets include individual-level HIV, HCV, and HSV2 status, stratified by age, sex, and ethnicity. Though descriptive analyses are ongoing, a statistical analysis across multiple surveys to estimate incidence of multiple diseases, including coinfection rates, has not been undertaken. Where cross-sectional, age-stratified prevalence studies are available, there is a range of methodologies which could be used to analyse them to estimate incidence. These are often used to estimate HIV incidence²³ and rarely used for HCV²⁴. Statistical methodologies range from catalytic models, classically applied to measles and other childhood diseases, and more recently applied to malaria, to complex transmission models. Simpler models have the advantage that they are easy to parameterise and understand, but can be lacking more detailed transmission mechanisms. In contrast, complex models include more detailed mechanisms and correlations but are correspondingly difficult to parameterise and analyse.

It is rare for data to include multiple time points, and statistical techniques become more difficult to apply when there are many parameters and changing transmission rates in different groups. Considering co-infections only amplifies this problem, and so it is clear that to understand the interactions between diseases a new approach is needed. To this end we developed a novel differential equation model and fitted it to the data from three JHH ED serosurveys using Markov chain Monte Carlo methods. We jointly modelled incidence of infection and coinfection with HIV, HCV and HSV2 within age/gender cohorts and then fitted the model to the observed prevalence within a non-parametric Bayesian framework, allowing us to infer incidence rates that vary smoothly with time and age. We estimated the incidence of HIV, HCV and HSV2 on coinfections of these viruses among the JHH ED population before predicting the burden of disease and testing our model predictions against the fourth and final serosurvey.

Methods

Data

At the adult JHH ED, identity-unlinked sero-surveys were conducted during six to eight-week periods in 2003, 2007, 2013 and 2016. These surveys are described in detail elsewhere^{21,25}. Briefly, excess sera were collected and assigned a unique study ID while chart review data were recorded in real-time. All laboratory testing was done after the collection period when the linked patient identifiers were removed from the dataset. Patient consent was waived by the ethical review board. This study was approved by the Johns Hopkins School of Medicine Institutional Review Board (IRB00083646, CIR00016268) and conducted by the ethical standards of the Helsinki Declaration of the World Medical Association. In the current analysis the authors had access only to anonymized data consisting of the test outcomes, sex, race (coded as black, white or other), age category (yearly from 18 to 89, and over 89) alongside the year the test was performed. The analysis excludes multiple results per unique subject, those < 18 years of age and individuals that did not self-identify as black or white, male or female. The reason for this limitation is due to sample size limitations. We also excluded 31 patients with incomplete data for one or more tests. Prevalence summaries of the data, showing the changing age-profile of infection, are provided in [Figure 1](#), and aggregated data is shown in [Table 2](#).

Figure 1. Prevalence of infection by disease, race, gender, age-group and survey year.

Demographic data / death rates

In order to accurately model the movement of individuals through the various infection (and co-infection) classes, we need to know the excess mortality due to infection with HCV or HIV. HIV-specific mortality rates are taken from 26, a study estimating the age-stratified mortality of HIV positive individuals on antiretroviral therapy (ART). These rates may underestimate the excess mortality in our population since: the rates are not race-stratified; and the study population are all on ART. Conversely, these rates may overestimate mortality since: individuals on ART are unlikely to be recent infections; the study includes individuals with HCV; and the rates are not described relative to the background death rate. Table 1 gives the age- and year-specific mortality rates taken from 26. We assumed that the mortality rates remained constant after 2008.

HCV-specific mortality rates were obtained from Mahajan *et al.*, a cohort-study of HCV- infected patients which reports an annual mortality rate of 12.854% in the cohort, compared with 1.046% in the general population²⁷. However to be included in the cohort, patients must be aware of their HCV status, which is not the case for the JHH ER data. Denniston *et al.* report that 50.3% of people detected with HCV were unaware they were infected²⁸. We assumed that infected individuals needed to be aware to be included in the cohort described in Mahajan *et al.* and, since patients that are aware are more likely to have experienced symptoms, we assume that unaware individuals experienced a death rate similar to the general population. This results in a HCV mortality rate of 5.868% annual excess death rate. Finally Thomas *et al.* find that end-stage liver disease is 3.67 times more likely in those over 38 years of age²⁹, and so we reduce the death rate from 5.868% for over 37s to 1.6% for under 38s.

Model

We developed a novel cohort model (illustrated in Figure 2), in which individuals are born with no infection, and then, over their lifetime may acquire each of the three infections, in any order. Each vertex of the cube in Figure 2 represents a state an individual may be in. Each state is denoted by three digits, giving the status of each different disease (0 if uninfected with that disease and 1 if infected). The digits are given in the order: HCV, HIV and HSV2 so that, for example, 101 is the state of being infected with HCV and HSV2 and uninfected with HIV. We use this notation to write down a system of eight ordinary differential equations (ODEs), tracking the proportion of the population in each state over time. We divide the population into “cohorts” of individuals of similar ages and for each cohort we use the same eight ODEs, but allow the parameters (the λ 's and θ 's) to be different for different cohorts. Each disease has a baseline infection rate for uninfected individuals (λ_1 , λ_2 and λ_3). The increased (or decreased) risk of infection due to previous infection status is given by the θ 's. For example, the relative risk of infection with HIV of an individual that is already infected with HCV (but not HSV2) is given by θ_1 . If θ_1 is greater than (less than) 1, we expect that prior infection with HCV increases (decreases) the risk of infection with HIV.

In addition to the events shown as arrows in Figure 2, we also include a variable death rate due to disease status. Since the natural death rate affects all disease states equally, without changing the proportions of the cohort

Table 1. Age- and time-stratified mortality rates for HIV positive individuals adapted from 26.

Data period	Age group	Mortality rate per 1000 person years
2003–2005	20–34	10.7
2003–2005	35–44	19.7
2003–2005	45–54	26
2003–2005	55+	30.8
2006–2008	20–34	11
2006–2008	35–44	12.5
2006–2008	45–54	18.1
2006–2008	55+	27.4

Figure 2. Model schematic. The schematic shows individuals moving between uninfected (0) and infected (1) for each of the infections: HCV; HIV; and HSV2, in that order. The rates of infection are represented as a varying force of infection, λ 's for each disease. Where individuals are already infected with at least one of the viruses, these forces of infection are modified by a factor, θ , which could increase or decrease risk.

in each state, we only need to consider the increase (or decrease) in death rate due to a particular disease status. For example, if infection with HIV increases the rate of death, then we would expect the proportions of the cohort in states 010, 110, 011 and 111 to decrease over time relative to the other states due to this.

We take p^{ijk} to be the proportion of the total population that have status i for HCV, j for HIV, and k for HSV2. So, for example, p^{011} is the proportion of the population that are negative for HCV and positive for both HIV and HSV2. Then our model is given by the following ordinary differential equations (ODEs):

$$\frac{\partial p^{000}}{\partial t} = (\delta_{\text{HCV}}p_{\text{HCV}} + \delta_{\text{HIV}}p_{\text{HIV}})p^{000} - (\lambda_1 + \lambda_2 + \lambda_3)p^{000} \quad (1)$$

$$\frac{\partial p^{100}}{\partial t} = (\delta_{\text{HCV}}p_{\text{HCV}} + \delta_{\text{HIV}}p_{\text{HIV}})p^{100} + \lambda_1p^{000} - (\theta_1\lambda_2 + \theta_2\lambda_3 + \delta_{\text{HCV}})p^{100} \quad (2)$$

$$\frac{\partial p^{010}}{\partial t} = (\delta_{\text{HCV}}p_{\text{HCV}} + \delta_{\text{HIV}}p_{\text{HIV}})p^{010} + \lambda_2p^{000} - (\theta_3\lambda_1 + \theta_4\lambda_3 + \delta_{\text{HCV}})p^{010} \quad (3)$$

$$\frac{\partial p^{001}}{\partial t} = (\delta_{\text{HCV}}p_{\text{HCV}} + \delta_{\text{HIV}}p_{\text{HIV}})p^{001} + \lambda_3p^{000} - (\theta_5\lambda_1 + \theta_6\lambda_2)p^{001} \quad (4)$$

$$\frac{\partial p^{110}}{\partial t} = (\delta_{\text{HCV}}p_{\text{HCV}} + \delta_{\text{HIV}}p_{\text{HIV}})p^{110} + \theta_1\lambda_2p^{100} + \theta_3\lambda_1p^{010} - (\theta_7\lambda_3 + \delta_{\text{HCV}} + \delta_{\text{HIV}})p^{110} \quad (5)$$

$$\frac{\partial p^{101}}{\partial t} = (\delta_{\text{HCV}}p_{\text{HCV}} + \delta_{\text{HIV}}p_{\text{HIV}})p^{101} + \theta_2\lambda_3p^{100} + \theta_5\lambda_1p^{001} - (\theta_8\lambda_2 + \delta_{\text{HCV}})p^{101} \quad (6)$$

$$\frac{\partial p^{011}}{\partial t} = (\delta_{\text{HCV}}p_{\text{HCV}} + \delta_{\text{HIV}}p_{\text{HIV}})p^{011} + \theta_4\lambda_3p^{010} + \theta_6\lambda_2p^{001} - (\theta_9\lambda_1 + \delta_{\text{HIV}})p^{011} \quad (7)$$

$$\frac{\partial p^{111}}{\partial t} = (\delta_{\text{HCV}}p_{\text{HCV}} + \delta_{\text{HIV}}p_{\text{HIV}})p^{111} + \theta_7\lambda_3p^{110} + \theta_8\lambda_2p^{101} - \theta_9\lambda_1p^{011} - (\delta_{\text{HCV}} + \delta_{\text{HIV}})p^{111} \quad (8)$$

where $p_{\text{HCV}} = \sum_{i=0}^1 \sum_{j=0}^1 p^{1ij}$ and $p_{\text{HIV}} = \sum_{i=0}^1 \sum_{j=0}^1 p^{i1j}$. Here δ_{HCV} and δ_{HIV} represent the additional risk of death due to infection with HCV and HIV, respectively.

Note that since we are considering proportions of the total population, the primary way that the additional risk of death is seen in the equations is through the reduction of total population size. This corresponds to an increase in each of the other proportions due to the reduced denominator.

Bayesian inference approach

The infection rates (incidences) were allowed to vary smoothly with time and age cohort. This flexible approach allows trends in incidence to be inferred from the data. The parameters are estimated in a Bayesian framework using an adaptive Markov chain Monte Carlo algorithm³⁰ and implemented in R³¹. The prior for the incidence parameters is a first order Gaussian random walk³², which implies that the difference between incidences at consecutive timepoints/cohorts is Normally distributed with mean zero. This is analogous to a Gaussian process³³ evaluated at discrete locations. The variances of the normal distributions describe the smoothness of the incidence surface, and are also estimated from the data. There are six smoothness parameters for each gender and race cohort, one for time and one for age for each of the three viruses. The nine relative risk parameters (the θ 's, see Figure 2) and the initial state of the model in 2003 were also estimated from the data.

We model the proportion of individuals with each infection status separately for each age cohort: birth year before 1943, and then each subsequent birth year up to 1995. We take a Bayesian non-parametric approach and estimate the three infection rates (incidences) for each year of the study (2003 to 2013) and for each age cohort. We therefore produce three incidence surfaces across time and age cohort. Because there is only a small amount of data for each cohort, we assume that the incidence surfaces are smooth, and so they can “borrow strength” from neighbouring points. Although estimating these incidence surfaces is the main challenge of the inference, in addition we must also estimate the 9 relative risk parameters ($\theta_1, \dots, \theta_9$), the initial conditions for each age cohort and the hyperparameters that describe the smoothness of the incidence surfaces.

More precisely, let $\lambda_{i,j,k}$ represent the rate of infection for disease i in year j for age cohort k , for $i \in \{1, 2, 3\}$, $k \in \mathbb{K}$ and $j \in \{2003, \dots, 2015\} \cap \mathbb{A}_k$, where $\mathbb{K}$ is the index set of age-cohorts and $\mathbb{A}_k$ is the set of study years after which cohort k reaches 18 years-old. The infection rates can be interpreted as the proportion of a completely susceptible population that would be infected in a year, and therefore we refer to them as incidences. Since we have data at the yearly resolution, we assume that the incidence is constant within each year.

Let $\mathbf{p}_k = (p_k^{000}, p_k^{001}, \dots, p_k^{111})$ represent the vector of initial conditions for age-cohort k , which must sum to 1. These vectors represent either the disease status at the start of the study in 2003; or, for the younger cohorts, the disease status when the cohort turns 18.

Priors. We assume a first order Gaussian random walk prior for the incidences across both time and age-cohort. For all $i \in \{1, 2, 3\}$,

$$\lambda_{i,j+1,k} - \lambda_{i,j,k} \sim N(\lambda_{i,j,k} - \lambda_{i,j-1,k}, \kappa_{i,\text{time}}^{-1}),$$

$$\lambda_{i,j,k+1} - \lambda_{i,j,k} \sim N(\lambda_{i,j,k} - \lambda_{i,j,k-1}, \kappa_{i,\text{age}}^{-1}).$$

This definition includes 6 hyperparameters (κ 's) that control the smoothness of the surfaces. We assumed that independent Gamma distributed priors with shape parameter 1 and rate parameter 0.01. Since this Gaussian random walk prior only specifies the differences between the lambdas, we also assumed that the mean level of each incidence surface follows an exponential distribution with mean 1%, to provide a mild shrinkage effect on the mean incidences.

The priors for the initial conditions were assumed to be Dirichlet distributions (which ensures that they sum to 1) with parameters equal to $\alpha = 0.8/8$, following the uninformative prior suggested by Berger *et al.*³⁴. The priors for the relative risk parameters $\theta_1, \dots, \theta_9$ were assumed to be independent exponential distributions with mean 1.

Likelihood. We assumed a multinomial likelihood function. Let $d_{j,k}^m$ be the number of individuals that tests indicate are in disease state $m \in \{(000), (001), \dots, (111)\}$ in year j from age-cohort k . Let $\mathbf{d}_{j,k}$ represent the data

vector for cohort k in year j and θ represent the complete vector of parameters. To evaluate the likelihood for cohort k in year j we start with the initial conditions $\mathbf{p}_k$ and solve the differential Equation (1)–Equation (8) forward in time until year j , call this $\mathbf{p}_k(j)$. This was done using the `rk4` Runge-Kutta differential equation solver from the R package `deSolve`³⁵. The log likelihood of observing data $\mathbf{d}_{j,k}$ given $\mathbf{p}_k(j)$ is simply $\sum_m d_{j,k}^m \log(p_k^m(j))$.

The complete log likelihood is then just the sum over all these component parts.

Markov chain Monte Carlo algorithm. To draw samples from the posterior distribution of the model we used Markov chain Monte Carlo. To update the $\lambda_{i,j,k}$ we used two kinds of single-site proposal. For j and k not on the boundaries we used conditional prior proposals³⁶ for 50% of the updates. For all remaining updates we used a Metropolis-Hastings Gaussian random walk, with proposal variance automatically tuned to 44% by rescaling by a factor x_n at iteration n when the iteration is accepted, and by $x_n^{0.44/(0.44-1)}$ when iteration n is rejected. To achieve diminishing adaptation, we used the sequence $x_n = 1 + 20/(20 + 0.2n)$. The precision parameters controlling the smoothness were updated using a Gibbs step.

If we consider only the data for the initial time point, the vector of initial conditions for cohort k , $\mathbf{p}_k$, has full conditional distribution $\mathbf{p}_k | \mathbf{d}_{1,k} \sim \text{Dir}(\alpha \mathbf{1} + \mathbf{d}_{1,k})$. For the final cohort there is no subsequent data and so the full conditional can be used as a Gibbs step. For earlier cohorts we cannot evaluate the posterior without first solving the system of differential equations. To find an efficient proposal we developed our own adaptive algorithm for Dirichlet distributions (similar to the algorithm in 37) that balances the above conditional distribution against the current location of the chain. Our proposal for $\mathbf{p}_k$ was $\mathbf{p}'_k \sim \text{Dir}(\alpha \mathbf{1} + \mathbf{d}_{1,k} + \beta_k \mathbf{p}_k)$, which was then accepted or rejected based on the usual Metropolis-Hastings ratio. To find an appropriate value for β_k for each cohort k , we applied the following adaptation algorithm during the burn-in phase of the MCMC. If a proposal was accepted then $\beta_k \mapsto \max\{0, \beta_k - 3\}$, and if a proposal was rejected then $\beta_k \mapsto \beta_k + 1$. This leads to an acceptance rate for $\mathbf{p}_k$ of roughly 25%, unless there is little data after the initial condition, in which case $\beta_k \approx 0$ and the acceptance rate is much higher.

The relative risk parameters $\theta_1, \dots, \theta_9$ were updated with single site Gaussian random walk proposals. To improve the mixing between these highly correlated parameters, we also added a joint update of the θ 's and the means of the 3 incidence surfaces using a Metropolis-Hastings Gaussian random walk in 12 dimensions, with the proposal covariance estimated adaptively using the approach of 30, implemented using the accelerated shaping algorithm³⁸.

Results

Incidence

Figure 3 shows the posterior median incidence rates in uninfected individuals for the three diseases as a function of time and age-cohort, with lower and upper confidence intervals shown in Figures S2, S3 and S4 (*Extended data*³⁹). For most groups and most infections we see a much stronger age-cohort effect than a time effect. That is, incidence in a cohort of individuals born in the same year remains approximately constant over time, while an overall decrease in incidence results from lower incidence in younger age groups. This effect is particularly prominent in HCV in black females, black males and white males and in HIV infection in black females (Figure 3). Incidence of all three infections is generally higher in the black population (the first and third rows in Figure 3) than the white population (rows two and four). In most graphs we see a small increase in incidence in the youngest age group in the most recent year. However, this is likely to be a statistical artefact resulting from increasing uncertainty due to fewer age groups (and thus less data) in more recent years. This increase in uncertainty when incidence rates are low can lead to an increase in the median, and so the uncertainty in the estimates is important in these regions (see Figures S1, S2 and S3 of the *Extended data* for confidence intervals³⁹).

The first column of Figure 3 shows HCV incidence. Black females, black males and white males all display a strong cohort effect, with individuals born around 1960 experiencing consistently high incidence, while younger cohorts show progressively lower incidence. The posterior median annual incidence peaks at 3.22% (90% CI: 1.66 – 4.76%) in 2003 for black males born in 1953. The peak incidence for black females was 2.95% (1.45 – 4.42%) for those born in 1957 and 2.63% (1.35 – 4.06%) for white males born in 1961, both also in the earliest year in the study. In contrast white females have consistently low incidence, decreasing in all cohorts to between 1.09 and 1.87% for the most recent year.

Figure 3. Incidence rates. Inferred incidence rates of Hepatitis C Virus (HCV), Human Immunodeficiency Virus (HIV) and Herpes Simplex Virus (HSV2) infection amongst black females (BF), white females (WF), black males (BM) and white males (WM) as a function of time and year of birth.

The second column of [Figure 3](#) shows HIV incidence. HIV incidence is more constant than HCV, and is not as smooth. Incidence in black females and white females is quite constant over time with posterior median incidences in the ranges 1.04–1.91% and 0.52–1.59%, respectively. However, the younger cohorts of black females appear to display somewhat reduced incidence. There is also some indication of reduced incidence in black males in both younger cohorts and more recent years. For black males born in 1962 the posterior median incidence drops from a peak of 1.94% at the start of the study to 1.40% in 2012; however, these declines are small compared to the level of noise (see [Figure S2](#) of the *Extended data* for HIV confidence intervals³⁹).

The third column of [Figure 3](#) shows posterior median HSV2 incidence. Black females show increasing incidence of HSV2 for later birth cohorts, peaking at 8.34% (90% CI: 5.09–11.5%) in the latest cohort, although there is some suggestion of decreasing risk over time for each cohort. This indicates a higher risk for younger individuals, but decreasing slightly as they age. In contrast, white females', black males' and white males' incidence of HSV2 does not change as much as for black females. White females also show increasing risk for later born cohorts, with a maximum of 3.82% (2.41–5.16%) for individuals born in 1978 and

then a decrease for the most recent cohorts. For each cohort, incidence decreases somewhat over time. A similar effect is seen in black males, with the highest-risk cohort born in 1973 with an incidence of 3.61% (1.99–5.26%) and general decreasing incidence over time for each cohort. Black males also display a low-risk cohort born around 1960, with incidence dropping to 1.81% (0.51–3.34%) for those born in 1958. Incidence in white males again shows decreasing risk over time for each cohort and relatively flat incidence between the cohorts.

Relative risks

Table 3 shows the posterior values of the θ parameters (see **Figure 2**). These give the relative risk of acquiring an infection for individuals that already have another infection. For example, θ_1 gives the relative risk of acquiring HIV for individuals that already have HCV. If $\theta_1 > 1$ (or $\theta_1 < 1$) then this would indicate that individuals with HCV are more likely (less likely, respectively) to acquire HIV than uninfected individuals.

Individuals that already have HSV2 are at lower risk of acquiring HIV or HCV (θ_5 and $\theta_6 < 1$). This has the possible explanation that individuals that are infected with HSV2 before any other disease are not therefore at high risk of acquiring HIV or HCV. Conversely, individuals that acquire HIV or HCV first are at greater

Table 2. Observed count of serological statuses from 2003, 2007 and 2013 surveys.

Serological status	Black female	Black male	White female	White male	Other (excluded)
000 - Sero-negative	926	1034	884	929	430
001 - HSV2	2236	903	408	210	155
010 - HIV	13	42	2	15	23
011 - HIV+HSV2	134	112	5	16	28
100 - HCV	53	241	39	143	2
101 - HCV+HSV2	322	352	96	82	8
110 - HCV+HIV	4	67	0	7	4
111 - HCV+HIV+HSV2	18	169	14	14	4
Total	3806	2920	1448	1416	685*

*including 31 with incomplete test data (see methods).

Table 3. Posterior medians and 90% credible intervals for the relative risk parameters (see **Figure 2 for interpretations).**

	Relative risk description	Black female	Black male	White female	White male
θ_1	HCV+, acquiring HIV	0.23 (0.02, 1.11)	0.27 (0.03, 0.83)	0.28 (0.02, 1.17)	0.19 (0.02, 0.76)
θ_2	HCV+, acquiring HSV2	5.25 (3.39, 8.04)	2.61 (1.65, 3.76)	5.35 (3.41, 7.72)	0.77 (0.10, 1.90)
θ_3	HIV+, acquiring HCV	0.40 (0.03, 1.82)	0.86 (0.07, 2.78)	1.01 (0.07, 4.22)	2.02 (0.32, 4.99)
θ_4	HIV+, acquiring HSV2	7.22 (4.63, 10.96)	7.51 (4.97, 10.59)	2.69 (0.76, 6.12)	4.32 (1.74, 7.83)
θ_5	HSV2+, acquiring HCV	0.10 (0.01, 0.26)	0.09 (0.01, 0.32)	0.10 (0.01, 0.38)	1.42 (0.51, 2.47)
θ_6	HSV2+, acquiring HIV	0.03 (0.00, 0.09)	0.12 (0.01, 0.38)	0.04 (0.00, 0.18)	0.15 (0.01, 0.64)
θ_7	HCV+ and HIV+, acquiring HSV2	0.91 (0.07, 3.19)	0.62 (0.06, 2.05)	1.73 (0.17, 5.90)	1.08 (0.09, 3.86)
θ_8	HCV+ and HSV2+, acquiring HIV	0.26 (0.02, 0.94)	0.20 (0.02, 0.73)	0.76 (0.10, 1.99)	0.49 (0.04, 1.54)
θ_9	HIV+ and HSV2+, acquiring HCV	2.34 (1.29, 3.65)	2.79 (1.56, 4.37)	1.41 (0.13, 5.21)	1.27 (0.15, 3.78)

risk of becoming infected with HSV2 (θ_2 and $\theta_4 > 1$). For those individuals eventually coinfecting with HIV and HSV2, if they are infected with HIV first, they rapidly become infected with HSV2. The parameter θ_7 represents a similar mechanism, however due to infrequent occurrence of coinfection with both HIV and HCV (due to high death rates), there is insufficient evidence to determine whether θ_7 is greater than or less than 1.

The other θ parameters are mostly inconclusive, but there is a suggestion of reduced infection with HIV when already infected with HCV ($\theta_1 < 1$), reduced infection with HIV when already infected with both HCV and HSV2 ($\theta_8 < 1$) and increased infection with HCV when already infected with both HIV and HSV2 ($\theta_9 > 1$).

Forward projections

We extended the random walk prior on the incidences to predict the incidences from 2014 to 2016, with associated uncertainty. The model equations were then solved from 2013 onwards to produce the posterior predictive distribution of prevalences in 2016 with the data from 2016 held-out for validation purposes.

Projections for 2016 were generally quite good for HCV and HIV (bottom rows of [Figure 4](#), [Figure 5](#), [Figure 6](#) and [Figure 7](#), whilst in general overestimating for HSV2. For white males and females the HIV prevalence was very low and the prevalence was overestimated in 2016.

Discussion

Incidence of infections, particularly stratified by key demographic variables, are an essential tool in developing effective control policies, planning for future treatment demand and targeting services. However, direct estimation of incidence of infections like HIV, HCV and HSV2 are almost impossible with current surveillance systems, and even tracking new diagnoses can hide underlying patterns due to changes in health-seeking behaviour in such long-lived infections. Therefore, methods for estimating incidence indirectly are an essential part of the toolbox of understanding these infections. Here we have developed a method of estimating incidence from a relatively cheap and infrequent survey design in a dynamic population of at-risk individuals.

By fitting our model within a Bayesian framework we have been able to fully characterise the uncertainty involved in this estimation procedure. We let the incidences vary smoothly in time and by age cohort, to enable us to borrow strength from similar data to reduce uncertainty, while at the same time providing distinct estimates for each age-cohort through time. This is illustrated in [Figure 4](#), where the uncertainty surrounding the prevalences is much smaller than the binomial CI surrounding the data, in which each observation is treated independently. We applied smoothing to the incidences but not to the initial prevalence distributions, which sometimes leads to noisy prevalence estimates (see [Figure 5](#) in particular). If smoothed estimates of prevalence are required a smoothing penalty prior could be used in place of the independent Dirichlet priors on the initial conditions.

Our study is of a population with a greater burden of HCV, HIV and HSV2 infection than the general population of the United States^{16,40–42}. Data on the prevalence and incidence of these infections in the general Baltimore City population are limited. The observed declines in HCV incidence in young black males and females are consistent with reported declines in HCV incidence in a predominantly-black cohort of community-based people who inject drugs in Baltimore⁴³. In addition, and consistent with our model-based estimates, the observed incidence of HCV between 2003 and 2016 was significantly higher in white ED patients as compared to black ED patients⁴⁴. The observed declines in HCV incidence in young black populations compared to white populations were also seen recently in New York City (NYC)⁴⁵. The extremely high incidence in HSV2 and/or HCV after HIV infection is also supported by another finding that shows a very small percentage of women in NYC are infected with HIV only⁴⁶.

A feature of our estimates is that incidence of disease at a particular age and time is heavily influenced by year of birth. One could speculate that this could indicate a strong impact on behaviour from early influences^{47,48}, or through the behaviour of individuals of a similar age who socialise together⁴⁹. This emphasises the potential importance of age-cohort counselling and early intervention with peers.

In addition to estimating incidence between surveys, we were able to forward-project from the 2013 survey to the 2016 data, to which the model was not fitted, with moderate success. Our predictions for HCV and HIV were generally good, but the predictions for HSV2 appeared to systematically overestimate the prevalence by a substantial margin. This large change in prevalence across all ages and ethnicities ([Figure 1](#)) is thought to be

Figure 4. Model fit and predictions for black females. Posterior median prevalence for black females (solid line) and 90% credible intervals (shaded) line for three diseases (columns) in the four sero-surveys (rows), each separated by age cohort. The model has been fitted to the observed data (crosses) in the first three rows of plots but data in the bottom row was held-out for validation purposes. Vertical lines represent 90% binomial CI for the observed data.

Figure 5. Model fit and predictions for white females. Posterior median prevalence for white females (solid line) and 90% credible intervals (shaded) line for three diseases (columns) in the four sero-surveys (rows), each separated by age cohort. The model has been fitted to the observed data (crosses) in the first three rows of plots but data in the bottom row was held-out for validation purposes. Vertical lines represent 90% binomial CI for the observed data.

Figure 6. Model fit and predictions for black males. Posterior median prevalence for black males (solid line) and 90% credible intervals (shaded) line for three diseases (columns) in the four sero-surveys (rows), each separated by age cohort. The model has been fitted to the observed data (crosses) in the first three rows of plots but data in the bottom row was held-out for validation purposes. Vertical lines represent 90% binomial CI for the observed data.

Figure 7. Model fit and predictions for white males. Posterior median prevalence for white males (solid line) and 90% credible intervals (shaded) line for three diseases (columns) in the four sero-surveys (rows), each separated by age cohort. The model has been fitted to the observed data (crosses) in the first three rows of plots but data in the bottom row was held-out for validation purposes. Vertical lines represent 90% binomial CI for the observed data.

due to changes in the surveyed population, for example due to the implementation of the Affordable Care Act (ObamaCare) in 2014, but this did not seem to have a strong effect on the prevalence of the other diseases. Whilst the results show that incidence rate trends were not constant over this three-year time-period, nevertheless the estimates would have provided useful predictions for public health policy. These forward projections would not have been possible from the prevalence data alone, but required the fitting of a model to the data to estimate incidence rates by age and sex.

The dynamics, incidence and prevalence of coinfection across diseases with similar or related pathways of transmission are an ongoing challenge for surveillance and control. In our analysis of these coinfection data we were able to estimate rates of individual infections after the initial infection; however, there are very few studies with the breadth of coinfection data which is present in this dataset, allowing such an analysis.

Strengths and limitations

The main strength of our approach was that we were able to infer cohort-specific incidence rates that varied through time, and therefore were able to identify trends in incidence at the age, gender and race cohort level. Our results showed that there could be large differences in incidences between age cohorts, which may not have been noticeable from a more naive analysis of the data.

A key weakness of our method was that for our incidence rates to be well-calibrated, we require an accurate estimate of the death rates in HIV and HCV infected populations, over and above the death rate in uninfected populations. This is because the death rates and infection rates are not both identifiable from the proportion of individuals in each infection state. The death rates proved challenging to obtain from the literature, particularly as HIV death rates may have improved substantially in recent years. Nonetheless, if these death rates are stable but incorrectly estimated, any qualitative trends in incidence informed from the data will remain correct.

Conclusions

We have developed a method for estimating age-specific incidence from anonymized cross-sectional prevalence surveys. This approach adds value to the data by providing age and time-specific incidence estimates which could not be obtained any other way and allows forecasting of future incidence. Our findings highlight a cohort-based trend that emphasizes the importance of age-cohort counselling and early intervention at a young age.

Data availability

Underlying data

Zenodo: drsimonspencer/HIV-HCV-HSV2-coinfection: Source code and extended data. <https://doi.org/10.5281/zenodo.5078271>³⁹.

This project contains the following underlying data:

- dataRANDOMIZED.csv
- data2016RANDOMIZED.csv

Extended data

Zenodo: drsimonspencer/HIV-HCV-HSV2-coinfection: Source code and extended data. <https://doi.org/10.5281/zenodo.5078271>³⁹.

This project contains the following extended data:

- ExtendedData.pdf

Data are available under the terms of the [Creative Commons Zero "No rights reserved" data waiver](#) (CC0 1.0 Public domain dedication).

Code availability

Source code available from: <https://github.com/drsimonspencer/HIV-HCV-HSV2-coinfection>

Archived source code at time of publication: <https://doi.org/10.5281/zenodo.5078271>³⁹

License: [GNU General Public Licence version 3](#)

References

- World Health Organization: **Global health observatory (GHO) data: HIV/AIDS global situation and trends 2018**. Accessed 28/08/19.
- Polaris Observatory HCV Collaborators: **Global prevalence and genotype distribution of hepatitis C virus infection in 2015: a modelling study**. *Lancet Gastroenterol Hepatol*. 2017; 2(3): 161–176.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Looker KJ, Magaret AS, Turner KME, *et al.*: **Global estimates of prevalent and incident herpes simplex virus type 2 infections in 2012**. *PLoS One*. 2015; 10(1): e114989.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Joint United Nations Programme on HIV/AIDS: **Fast-track: ending the aids epidemic by 2030**. 2014; Accessed 28/08/19.
- World Health Organization: **Combating hepatitis B and C to reach elimination by 2030: advocacy brief**. 2016; Accessed 28/08/19.
[Reference Source](#)
- Landsberger HA: **Hawthorne Revisited: A Plea for an Open City**. Cornell University, 1957.
- Kassanjee R, Pilcher CD, Keating SM, *et al.*: **Independent assessment of candidate HIV incidence assays on specimens in the CEPHIA repository**. *AIDS*. 2014; 28(16): 2439–49.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Laeyendecker O, Konikoff J, Morrison DE, *et al.*: **Identification and validation of a multi-assay algorithm for cross-sectional HIV incidence estimation in populations with subtype C infection**. *J Int AIDS Soc*. 2018; 21(2): e25082.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Shepherd SJ, Kean J, Hutchinson SJ, *et al.*: **A hepatitis C avidity test for determining recent and past infections in both plasma and dried blood spots**. *J Clin Virol*. 2013; 57(1): 29–35.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Patel EU, Cox AL, Mehta SH, *et al.*: **Use of hepatitis C virus (HCV) immunoglobulin G antibody avidity as a biomarker to estimate the population-level incidence of HCV infection**. *J Infect Dis*. 2016; 214(3): 344–352.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Morrow RA, Friedrich D, Krantz E, *et al.*: **Development and use of a type-specific antibody avidity test based on herpes simplex virus type 2 glycoprotein G**. *Sex Transm Dis*. 2004; 31(8): 508–515.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Justman JE, Mugurungi O, El-Sadr WM: **HIV Population Surveys - Bringing Precision to the Global Response**. *N Engl J Med*. 2018; 378(20): 1859–1861.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Coates TJ, Kulich M, Celentano DD, *et al.*: **Effect of community-based voluntary counselling and testing on HIV incidence and social and behavioural outcomes (NIMH Project Accept; HPTN 043): a cluster-randomised trial**. *Lancet Glob Health*. 2014; 2(5): e267–77.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Solomon SS, Mehta SH, McFall AM, *et al.*: **Community viral load, antiretroviral therapy coverage, and HIV incidence in India: a cross-sectional, comparative study**. *Lancet HIV*. 2016; 3(4): e183–90.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Chemaitelly H, Nagelkerke N, Omori R, *et al.*: **Characterizing herpes simplex virus type 1 and type 2 seroprevalence declines and epidemiological association in the United States**. *PLoS One*. 2019; 14(6): e0214151.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Rosenberg ES, Rosenthal EM, Hall EW, *et al.*: **Prevalence of Hepatitis C Virus Infection in US States and the District of Columbia, 2013 to 2016**. *JAMA Netw Open*. 2018; 1(8): e186371.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Kelen GD, Fritz S, Qaqish B, *et al.*: **Substantial increase in human immunodeficiency virus (HIV-1) infection in critically ill emergency patients: 1986 and 1987 compared**. *Ann Emerg Med*. 1989; 18(4): 378–82.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Kelen GD, Hexter DA, Hansen KN, *et al.*: **Trends in human immunodeficiency virus (HIV) infection among a patient population of an inner-city emergency department: implications for emergency department-based screening programs for HIV infection**. *Clin Infect Dis*. 1995; 21(4): 867–75.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Patel EU, Laeyendecker O, Hsieh YH, *et al.*: **Parallel declines in HIV and hepatitis C virus prevalence, but not in herpes simplex virus type 2 infection: A 10-year, serial cross-sectional study in an inner-city emergency department**. *J Clin Virol*. 2016; 80: 93–7.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Patel EU, Frank MA, Hsieh YH, *et al.*: **Prevalence and factors associated with herpes simplex virus type 2 infection in patients attending a Baltimore City emergency department**. *PLoS One*. 2014; 9(7): e102422.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Kelen GD, Hsieh YH, Rothman RE, *et al.*: **Improvements in the continuum of hiv care in an inner-city emergency department**. *AIDS*. 2016; 30(1): 113–20.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Hsieh YH, Rothman RE, Laeyendecker OB, *et al.*: **Evaluation of the Centers for Disease Control and Prevention recommendations for hepatitis C virus testing in an urban emergency department**. *Clin Infect Dis*. 2016; 62(9): 1059–1065.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Hallett TB, Zaba B, Todd J, *et al.*: **Estimating incidence from prevalence in generalised HIV epidemics: methods and validation**. *PLoS Med*. 2008; 5(4): e80.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Lehman EM, Wilson ML: **Epidemic hepatitis C virus infection in Egypt: estimates of past incidence and future morbidity and mortality**. *J Viral Hepat*. 2009; 16(9): 650–658.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Kelen GD, Fritz S, Qaqish B, *et al.*: **Unrecognized human immunodeficiency virus infection in emergency department patients**. *N Engl J Med*. 1988; 318(25): 1645–1650.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Samji H, Cescon A, Hogg RS, *et al.*: **Closing the gap: increases in life expectancy among treated hiv-positive individuals in the United States and Canada**. *PLoS One*. 2013; 8(12): e81355.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Mahajan R, Xing J, Liu SJ, *et al.*: **Mortality among persons in care with hepatitis C virus infection: the Chronic Hepatitis Cohort Study (CHACS), 2006–2010**. *Clin Infect Dis*. 2014; 58(8): 1055–1061.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Denniston MM, Kleven RM, McQuillan GM, *et al.*: **Awareness of infection, knowledge of hepatitis C, and medical follow-up among individuals testing positive for hepatitis C: National health and nutrition examination survey 2001–2008**. *Hepatology*. 2012; 55(6): 1652–1661.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Thomas DL, Astemborski J, Rai RM, *et al.*: **The natural history of hepatitis C virus infection: host, viral, and environmental factors**. *JAMA*. 2000; 284(4): 450–456.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Haario H, Saksman E, Tamminen J: **An adaptive Metropolis algorithm**. *Bernoulli*. 2001; 7(2): 223–242.
[Publisher Full Text](#)
- R Core Team: **R: A Language and Environment for Statistical Computing**. R Foundation for Statistical Computing, Vienna, Austria, 2018.
- Fahrmeir L, Lang S: **Bayesian inference for generalized additive mixed models based on Markov random field priors**. *J R Stat Soc Ser C Appl Stat*. 2001; 50(2): 201–220.
[Publisher Full Text](#)
- Williams CKI, Rasmussen CE: **Gaussian processes for machine learning**. the MIT Press, 2006; 2(3): 4.
[Reference Source](#)
- Berger JO, Bernardo JM, Sun D: **Overall objective priors**. *Bayesian Anal*. 2015; 10(1): 189–221.
[Publisher Full Text](#)
- Soetaert K, Petzoldt T, Setzer RW: **Solving Differential Equations in R: Package deSolve**. *J Stat Softw*. 2010; 33(9): 1–25.
[Publisher Full Text](#)
- Knorr-Held L: **Conditional prior proposals in dynamic models**. *Scand Stat*. 1999; 26(1): 129–144.
[Publisher Full Text](#)
- Benschop J, Nisa S, Spencer SEF: **Still 'dairy farm fever'? A Bayesian model for leptospirosis notification data in New Zealand**. *J R Soc Interface*. 2021; 18(175): 20200964.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Spencer SEF: **Accelerating adaptation in the adaptive**

- Metropolis Hastings random walk algorithm.** (in revision), 2021.
39. Spencer S: **drsimonspencer/HIV-HCV-HSV2-coinfection: Source code and extended data.** July 2021.
[Publisher Full Text](#)
  40. McQuillan GM, Kruszon-Moran D, Flagg EW, *et al.*: **Prevalence of herpes simplex virus type 1 and type 2 in persons aged 14-49: United States, 2015-2016.** *NCHS Data Brief*. US Department of Health and Human Services, Centers for Disease Control and Prevention, 2018; (304): 1-8.
[PubMed Abstract](#)
  41. Patel EU, Gaydos CA, Packman ZR, *et al.*: **Prevalence and correlates of *trichomonas vaginalis* infection among men and women in the United States.** *Clin Infect Dis*. 2018; **67**(2): 211-217.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  42. Hess KL, Johnson SD, Hu X, *et al.*: **Diagnoses of HIV infection in the United States and dependent areas, 2017.** Technical report, Center for Disease Control, 2018.
  43. Mehta SH, Astemborski J, Kirk GD, *et al.*: **Changes in blood-borne infection risk among injection drug users.** *J Infect Dis*. 2011; **203**(5): 587-594.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  44. Hsieh YH, Patel AV, Loevinsohn GS, *et al.*: **Emergency departments at the crossroads of intersecting epidemics (HIV, HCV, injection drug use and opioid overdose)-estimating HCV incidence in an urban emergency department population.** *J Viral Hepat*. 2018; **25**(11): 1397-1400.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  45. Suryaprasad AG, White JZ, 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**(10): 1411-1419.
[PubMed Abstract](#) | [Publisher Full Text](#)
  46. Des Jarlais DC, Arasteh K, McKnight C, *et al.*: **HSV-2 co-infection as a driver of HIV transmission among heterosexual non-injecting drug users in New York City.** *PLoS One*. 2014; **9**(1): e87993.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  47. Ryser MD, Rositch A, Gravitt PE: **Modeling of Us human papillomavirus (HPV) seroprevalence by age and sexual behavior indicates an increasing trend of HPV infection following the sexual revolution.** *J Infect Dis*. 2017; **216**(5): 604-611.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  48. Liu G, Hariri S, Bradley H, *et al.*: **Trends and patterns of sexual behaviors among adolescents and adults aged 14 to 59 years, United States.** *Sex Transm Dis*. 2015; **42**(1): 20-26.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
  49. Mossong J, Hens N, Jit M, *et al.*: **Social contacts and mixing patterns relevant to the spread of infectious diseases.** *PLoS Med*. 2008; **5**(3): e74.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)