A quasi-markov model for transmission and disease elimination: Hepatitis C among people who inject drugs

Rachel Hart-Malloy† and Gregory DiRienzo‡*

*Correspondence: adirienzo@albany.edu
†These authors contributed equally to this work.
‡AIDS Institute, New York State Department of Health, Albany NY; Department of Epidemiology and Biostatistics, University at Albany, State University of New York, Rensselaer, USA.

Abstract

Background: The use of mathematic modeling to better understand the spread of Hepatitis C and the impact of interventions can be invaluable for localities, states and countries with a large burden of injection drug use. New York State (NYS) is estimated to be home to a large number of people who inject drugs (PWID), however the burden of hepatitis C among this population and the impact of interventions are not known and/or fully understood. Since accurate modeling can be complex and require costly data analysis software to implement, the purpose of this paper was to derive a methodology to accurately model the prevalence of hepatitis C at the state level that is tractable and easily implemented with free software.

Methods: A non-stationary quasi-Markov model, is proposed that is implemented using free software R©. The methodology aims to estimate hepatitis C prevalence and evaluate impacts of interventions on disease burden among PWID in NYS. The approach is “quasi-Markov” because transition probabilities among states change in time and depend on past information. Interventions evaluated aimed to: reduce sharing needles and/or other drug use paraphernalia; increase needle disinfection rates; and increase availability of clean needles and other drug use paraphernalia.

Results: The quasi-Markov model estimated hepatitis C prevalence among PWID reached an equilibrium value of 63.6 percent after 50 years. In order to eliminate disease, all proposed interventions were needed, resulting in an estimated prevalence less than 1.0 percent, 56 years after implementation. Using parameters defined for an alternate study modeling hepatitis C prevalence found similar results, serving to reinforce the validity of the proposed methods.

Conclusion: The results of this analysis demonstrate the feasibility of using a non-stationary quasi-Markov methodology to model the spread of hepatitis C among PWID using free programming software. Coding provided can be used by other researchers to use and modify for their own purposes and could further impact this field of research as well as inform and promote additional education prevention and interventions for PWID.

Keywords: Hepatitis C virus, markov modeling, injection drug use, people who inject drugs

Introduction

Hepatitis C virus (HCV) infection, estimated to be the most common chronic blood borne infection in the United States (US), has been shown to disproportionately affect people who inject drugs (PWID) [1-3]. Whereas prevalence estimates among the general US population ranges from 1.0 to 1.6 percent, prevalence estimates among PWID range from 40 to 90 percent [1-8]. HCV transmission among PWID can result from, but not limited to, sharing needles or other drug use paraphernalia (i.e., cookers, cotton, or rinse water; referred to as equipment) [9-11]. Individuals may be asymptomatic for decades and if left untreated, HCV infection can lead to cirrhosis of the liver, liver failure, hepatocellular carcinoma and death [12,13]. PWID, in particular, are at an even greater risk of the morbidity and mortality associated with untreated HCV because they often times do not have access to the healthcare system [3,8,14,15]. HCV prevention and interventions targeted towards PWID have included behavioral interventions, substance use treatment,
Methods

Non-stationary quasi-Markov modeling was utilized to estimate HCV prevalence and to assess the impact of interventions on disease elimination among PWID in NYS. The structure of the model was motivated from those results of Corson, et al., 2012, who utilized a deterministic compartmental model for HCV transmission [19]. Although the results of Corson, et al., 2012 are very valuable, their approach is somewhat sophisticated by way of use of deterministic differential equations that require high powered computing to implement, both of which may present a barrier to general researchers in this field for a thorough understanding and tractable implementation. Similar to Corson, et al., we defined a model whereby individuals transition between a set of possible predefined states through time. However, our approach uses a tractable and somewhat simple stochastic framework based on Markov modeling for governing transitions between states and is easily implemented with free software (Supplement Data). Possible routes of HCV infection modeled were through needle and equipment sharing only. The natural history of HCV was developed based upon previous literature and includes basic states of the disease which will be described in more detail below (Figure 1) [20]. The transition probabilities between states of HCV were dependent upon virologic parameters of the disease, such as the probability of spontaneously clearing the infection, and also parameters determined based upon characteristics of the population, such as the prevalence of HCV among PWID. For this reason, transition probabilities between states were non-stationary, meaning they were updated at each time unit to accommodate the changing parameters of the PWID population. The time unit chosen for this analysis was days, which we think realistically accommodates the real world behaviors in this population, and estimates were modeled over 75 years (27,375 days), which in our opinion is an acceptable time range based on expected human life span. Time zero is a theoretical year in which the PWID population is defined by the assumptions of the model, explained in more detail below, and is not intended to reflect a specific calendar year.

Quasi-markov model transition probabilities and parameters

A graphic showing the possible states of the progression of HCV and the allowable transitions between states is provided in Figure 1. The probability that PWID transition from susceptible state (s) (i.e., never infected with HCV) to acutely infected (state h or h) was based upon the probability of infection at time t, p(t) (defined as risk of infection due to sharing needles or equipment) [19]. In general terms, the probability of infection at time t depends upon a combination of the probabilities that an individual shares needles or equipment, the shared needle is not cleaned, the un-cleaned needle or equipment shared is infected, and lastly, the infected, shared and un-cleaned needle or equipment successfully transmits the disease during that sharing encounter. These probabilities are inherently dependent upon the proportion of the population that is acutely infected or chronically infected (state y) as well as the proportion of needles and equipment available. More specifically, the probability of infection experienced by the susceptible population is driven by the probability per day that the individual shares needles (λN) or equipment (λE), the probability that they do not clean needles (1-ϕN), the probability that the specific needle or equipment used will transmit infection successfully (ψN,ψE) and the proportion of infected needles (βNS(t)) and equipment (βES(t)) in the PWID population at time t, where s can be either h, h, or y and (ψN,ψE) refers to the probability of successful acute transmission when sharing needles and equipment, respectively (Table 1).

\[
p(t) = λN (1-ϕN) (ψN,h1 (t) + βNS (t)) + ψN,ψE (t) + λE (ψE,h2 (t) + βES (t))\]

The proportion of infected needles was dependent upon the
average number of injections per day (2.8) [37] and the average number of needles in NYS: 1) dispensed through syringe exchange programs (SEPs) (from 2002-2011); and 2) sold through the expanded syringe access program (ESAP) (Table 1) [30,38]. See Corson et al., for derivation of the proportion of infected needles [19]. The proportion of infected equipment was dependent upon the amount of equipment dispensed from SEPs. For the purpose of this analysis, equipment was considered to include cotton, cookers and rinse water. A determination of this quantity is not available. Therefore, the ratio of the amount of equipment available to the number of needles dispensed or sold was assumed to be 3:4 based on feedback from SEPs. This was varied in a sensitivity analysis, the details of which are described below.

Not all persons who become acutely infected develop...
chronic HCV infection; a proportion of acutely infected persons spontaneously clear the virus within six months [39]. Therefore, based on previous work, two acute HCV states \((h_1, h_2)\) were included in the natural history model for which acutely infected individuals in state \(h_1\) may progress to chronic HCV infection (state \(y\)) and acutely infected individuals in state \(h_2\) may clear the infection and progress to a susceptible state of previously infected individuals \((x)\) [19]. Given the time unit of days, new cases remain in state \(h_1, h_2\) for 180 days (six months). Only upon reaching day 180 does the model permit transition out of those states and into either state \(y\) or state \(x\), thus creating a transition probability that is dependent upon ones state in the previous time iteration. That is, for \(j=1, 2\), let denote the number of persons in state at day \(t=0, \ldots, 27,375\), where \(t=0\) is the initial state and \(\mu\) the rate individuals leave the PWID population (Table 1). Then, for \(t=1, \ldots, 180\):

\[
h_j(t) = \sum_{i=0}^{t-1} (h_j(v) - \mu)
\]

and for \(k=1, \ldots (27,375-180)\):

\[
h_j(180 + k) = \sum_{i=0}^{180-k-1} (h_j(v) - \mu) - \sum_{i=0}^{t-1} (h_j(v) - \mu)
\]

Some research has found that following spontaneous clearance of an acute infection a proportion of individuals may have protective immunity to re-infection; however, findings regarding a conferring of a protective immunity are equivocal [40-43]. In order to capture the possibility of some level of a protective immunity, a corresponding state \((z)\) was included in the natural history model with a low transition probability (0.25) [19].

Inherent in the model at each time unit is the probability of leaving the PWID population \((\mu)\), represented by the state exit \(e\). This transition is possible from all states of the natural history (including during the first 180 days of an incident case). Research has shown that the probability of an individual leaving the population is roughly equivalent to the probability of new individuals entering the population [31,44]. Thus, the proportion of the population that transitions to state \(e\) is set to equal the proportion of the population added into susceptible state, \(x\), at the end of each time iteration (a 1-to-1 replacement of those who left with those who enter). This decision was made solely for modeling purposes and does not reflect an individual leaving the population (regardless of infection status) and re-entering as susceptible. The parameters and the sources used to define the annual transition probabilities are in Table 1.

**Analysis**

The quasi-Markov model in Figure 1 was utilized to model HCV prevalence among PWID through needle and equipment sharing (model 1). To determine what interventions would be needed to eliminate HCV from the population, modifications using a basic reproductive number \((R_0)\) were explored and the initial model was altered to determine the effects different interventions would have on HCV prevalence over time (model 2).

**Modeling HCV prevalence: model 1**

Using the defined probabilities, the population distribution of PWID in NYS at time \(t=0\) was assumed to be 99.0 percent susceptible \((x)\) and 1.0 percent acutely infected only \((h_2)\). This is to reflect an assumption that only a small proportion of the population is infected at time \(t=0\). The probability that an individual shares needles and/or equipment per day was allowed to vary in the model based upon the proportion of the PWID population considered to be high (0.38), medium (0.19) or low risk (0.43) [31]. This distribution of PWID risk groups was informed by their needle and equipment sharing behaviors, as determined from the literature (Table 1) [31]. These probabilities were assumed to be held constant over time. Using the R© code developed for this analysis (Supplement Data), the population was projected through 75 years (27,375 time units). At each time point, the updated prevalence in each state was calculated as:

\[
M(t) = \begin{bmatrix} x(t+1), h_1(t+1), h_2(t+1), x_1(t+1), x_2(t+1), y(t+1), x_3(t+1) \end{bmatrix} \]

Where the transition matrix \(M(t)\) is defined in Figure 1.

HCV prevalence among the individuals was defined as those who have ever been infected, or those who would be defined as HCV antibody positive (anti-HCV), therefore summing the prevalence of individuals in the following states: \(h_1, h_2, x_1, y, z\), and \(z\).

Several sensitivity analyses were conducted. First, the ratio of available equipment to needles dispensed or sold in NYS was varied from 1:1 to 1:10. Second, rather than assume 99.9 percent of the population was susceptible \((x)\) at time \(t=0\), the proportion susceptible \((x)\) was varied from as low as 70.0 to as high as 99.9 to determine any differences in the prevalence once reaching equilibrium. Third, given the equivocal nature of protective immunity, the model was evaluated removing this state in the natural history. Lastly, the model was run assuming the proportion of individuals who spontaneously clear the infection (probability of transitioning from susceptible \((x, x_1)\) to acutely infected \((h_2)\)) was larger based upon results presented in the literature (0.26 compared to 0.23) [42].

**Modeling disease elimination through modifying parameters: model 2**

Modifiable parameters assessed were: 1) reducing the amount of needle and equipment sharing \((\lambda_w, \lambda_e)\); 2) increasing the proportion of PWID who properly sterilize needles \((\phi)\); 3) increasing the number of needles dispensed or sold in a given year and; 4) increasing the amount of equipment dispensed in a given year. To determine to what degree specific parameters
would need to be modified to eliminate HCV, defined here as a prevalence of less than 1%, a formula for the basic reproductive number ($R_0$) was used. The number $R_0$ is defined as number of secondary infections resulting from an infected individual. Briefly, when $R_0=1$, each infected individual is responsible for infecting another individual; therefore, when $R_0<1$, a disease will theoretically die out and when $R_0>1$, a disease will reach an epidemic level. The formula used for our purposes was based on a version of a previously derived equation for $R_0$ determined by Corson, et al., which was modified to include information on both needle and equipment sharing [19]. In brief, the formula derived by Corson, et al., defines $R_0$ as the expected number of infections resulting from an infectious needle multiplied by the expected number of needles generated by a PWID during their infectious lifetime. For our purposes, using the parameters defined for this study, $R_0$ was modified to incorporate the expected number of infections resulting from infectious equipment multiplied by the expected number of equipment (considered as cotton, cooker and water, as defined above) generated by a PWID during their infectious lifetime:

$$R_0 = \frac{\left(\frac{\lambda}{\mu} + 365.25)(1-\phi)\right)}{\left(\frac{\mu}{\mu} + 365.25\right)[(\frac{\mu}{\mu} + 365.25) + \sigma \left(1 + (\frac{r}{\mu})^{365.25}\right)]}$$

$$+ \frac{\left(\mu + 365.25\right)}{\left(\frac{\mu}{\mu} + 365.25\right)[(\frac{\mu}{\mu} + 365.25) + \sigma \left(1 + (\frac{t}{\mu})^{365.25}\right)]}$$

where $1/\sigma=0.5$, as defined by Corson, et al., represents the duration of an acute phase in years [19].

The number $R_0$ was calculated for low, medium and high risk PWID using the same behavioral and virologic parameters utilized in model 1. Assuming $R_0$ was greater than one (no disease elimination), it was determined how to adjust the parameters representative of the above described interventions in order to ensure that $R_0$ would be less than one (allowing disease elimination), thus creating the theoretical conditions necessary to eliminate HCV. Adjustments were made using an empirical trial and error process.

The parameters were then updated in model 1 to determine what prevalence would be over time given the interventions (model 2). The distribution of the population of PWID at time $t=0$ was defined based upon the resulting distribution from model 1. This population was then projected again through 75 years using the updated parameters. A sensitivity analysis was employed in which the ratio of available equipment to needles distributed or sold was varied from 1:1 to 1:10.

**Results**

**HCV prevalence**

Estimated prevalence among PWID in NYS from model 1 rose to over 50.0 percent within 25 years and reached equilibrium at 64.1 percent within 50 years (Figure 2). Varying the ratio of equipment to needles dispensed or sold from the original 3:4 to 1:1 and 1:10, respectively, resulted in prevalence ranging from 63.0 to 67.6 percent (results not shown). Prevalence was not impacted when the proportion susceptible, x, at time $t=0$ was varied. When removing the transition probability for protective immunity, prevalence increased at a more rapid rate and reached a higher prevalence at equilibrium (65.8 percent) (Figure 2). When assuming the proportion that spontaneously clears the infection was higher, prevalence increased at a slower rate and the prevalence at equilibrium was lower at 63.3 percent (Figure 2).

**Disease elimination**

The basic reproductive number exceeded 1 for only the medium (4.3) and high risk groups (7.3). These parameters, referred to collectively as a multi-component approach, included modifying the parameters as follows: 1) reducing the probability of sharing needles in the high risk group by half (from a probability of 0.22 per day to 0.11 per day; the probability in the other risk groups was 0), 2) reducing the probability of sharing equipment in the high and medium risk groups by a quarter (from 0.40 and 0.51 to 0.10 and 0.13, respectively), 3) increasing the probability of cleaning ones needle in the high risk group by two-fold (from 0.145 to 0.290), 4) increasing the amount of equipment in the high and medium risk groups available by four-fold (from 3,812,583 to 15,250,332) and, 5) increasing the number of needles dispensed in the high and medium risk groups two and a half times (from 5,083,444 to 12,708,611). When prevalence was modeled assuming the above mentioned multi-component approach, the prevalence declined reaching 50.0 percent in less than 5 years (Figure 3). Upon reaching equilibrium, the prevalence was 0.2 percent. A visual display of the impact of each modified parameter is displayed in Figure 3.

Varying the ratio of available equipment to dispensed needles impacted what is required to eliminate disease and reach equilibrium. As would be expected, when assuming a 1:1 ratio, less equipment was needed to reach equilibrium (a three-fold and two-fold increase among the high and medium
The prevalence of HCV among PWID in NYS is unknown and needed a six-fold for the high risk group compared to the Scotland using a deterministic compartmental model using Corson, et al., did not account for equipment sharing. Contrary to our proposed methodology, the model used by Corson, et al., where equipment sharing was not accounted for. In addition, since Corson, et al., utilized shared injections per year in their model, the parameters were adjusted to reflect to the probability of sharing per day for our proposed methodology.

The model used by Corson, et al., determined that HCV prevalence among PWID in Glasgow rose steadily for the four-fold increase for the initial analysis). When assuming a 1:10 ratio, to reach equilibrium, adjustments were needed for the high risk group only requiring an increase in both the amount of needles (needed to increase four-fold for the high risk group compared to 2.5 times for the initial analysis) and equipment (needed a six-fold for the high risk group compared to the four-fold increase for the initial analysis) and a reduction in the probability of sharing needles per day (reduced by one-third compared to one-half for the initial analysis).

**Model validation**

The prevalence of HCV among PWID in NYS is unknown and therefore cannot be used to validate the model. However, in an effort to validate the findings, our model was run on previously published work modeling prevalence of HCV [19]. Here, Corson, et al., modeled HCV prevalence in Glasgow, Scotland using a deterministic compartmental model using the dynamic systems modeling software Berkeley Madonna™. Contrary to our proposed methodology, the model used by Corson, et al., did not account for equipment sharing.

Therefore, for comparison purposes, prevalence was modeled by our proposed methods using the parameter values implemented by Corson, et al., where equipment sharing was not accounted for. In addition, since Corson, et al., utilized shared injections per year in their model, the parameters were adjusted to reflect to the probability of sharing per day for our proposed methodology.

The model used by Corson, et al., determined that HCV prevalence among PWID in Glasgow rose steadily for the first decade reaching equilibrium around year 15 at an HCV prevalence of 68.9 percent. The resulting HCV prevalence results from our model using their parameters were similar with HCV prevalence rising steadily and reaching equilibrium later at year 35 at a slightly lower HCV prevalence of 62.1 percent. Although a similar prevalence was determined in Glasgow by Corson et al., as in NYS, the populations examined differed in several ways: the estimated number of PWID, the estimated number of needles distributed, the proportion of the population estimated to be sharing, and the probability of acute and chronic transmission per shared injection event. These differences thus resulted in differing PWID to needle ratios and needle turnover rate as well.

Modeling HCV prevalence when \( R_0 < 1 \), was also done in an effort to provide some validation. Similar to results found by Corson et al., elimination was reached using the parameters necessary to meet \( R_0 < 1 \). However, the decline in prevalence was slower with our model reaching <0.002 within in 35 years compared to 20 years as determined by Corson, et al.

In an effort to further validate the model, the estimate for HCV prevalence among PWID in NYS from this model was compared to other estimates among PWID. The prevalence of HCV among PWID has been cited to range between 40 and 90 percent, depending upon geography [8]. In addition, studies were done in the late 1990’s and early 2000’s in areas of New York City finding HCV prevalence ranging from 32-90 percent. The prevalence found in this study falls within the range of previous estimates [24,45,46].

Furthermore, the estimated number of secondary infections resulting from PWID belonging to the medium and high risk groups (4.3 and 7.3, respectively) is similar to other findings by Magiorkinis et al., (2013) who calculated \( R_0 \) in Greece for HCV subtypes using epidemiologic (surveillance) data and phylodynamic modeling [47]. For subtypes with a higher proportion of PWID, they estimated that the number of secondary infections were 3.4, 11.5 and 2.4 for subtypes 1a, 3a and 4a, respectively; the subtype with the highest proportion of PWID (47 percent) was subtype 3a [47]. The results from the present analysis thus fall within the range found by Magiorkinis et al., despite the differing methodologies.

Finally, two pathological cases were explored in an effort to validate the model in which the transmission probabilities were set to the unrealistic parameters of zero and 1. With the parameters set to zero, the steady-state prevalence went to zero; alternatively, with the parameters set to 1, likewise the steady-state prevalence went to 1 (i.e., all individuals would be infected).

**Discussion**

This study demonstrated that HCV prevalence can be estimated at the state level using a non-stationary quasi-Markov model run entirely using free software. To the best of our knowledge, this is the first study using a quasi-Markov model to estimate HCV prevalence. In addition, the other available methodologies
presented in the literature to estimate HCV prevalence require complex statistical analyses software packages that can be expensive. Although the model cannot be directly compared to existing NYS HCV prevalence among PWID due to lack of observed data, the results estimated an HCV prevalence within the range of other studies conducted in NYS and elsewhere among samples of PWID further supporting the validity of the method. In addition, results from this methodology demonstrated that using a multi-component intervention model was needed to effectively eliminate HCV from the population. This supports other research on the effectiveness of such harm reduction strategies among PWID.

Although it is beyond the scope of the paper to assess or recommend interventions that would result in the modifications as specified to eliminate HCV among PWID in NYS, the parameters were intentionally chosen as “modifiable” in nature.

It must be noted that time in this analysis does not reflect prevalence over time in NYS among PWID. This model is used as a tool to see what prevalence could be assuming a small proportion of PWID are initially infected, accounting for current IDU behaviors, and needle and equipment distribution.

The actual prevalence among PWID in NYS over time has likely been impacted by interventions implemented such as needle exchange and sale through SEPs and the ESAP, and the distribution of equipment by SEPs.

There are some limitations to this analysis. First, this study does not account for the possibility of dual infection, in which a person is infected with more than one HCV strain. This could impact the probability of spontaneous clearance. Further analysis examining what impact this would have on disease spread and elimination are needed. Second, model 1, though taking into account other current interventions such as needle and equipment distribution, does not account for the possibility of HCV treatment. The proportion of active PWID who are treated is likely low, given low overall treatment in the general population, and would therefore only marginally lower the overall prevalence. Third, parameters from localities other than NYS were used in this analysis. Although several studies conducted in NYS among PWID have been done, they have been restricted to areas in NY City and do not have the detailed level of information needed (daily probability of sharing needles and equipment, daily probability of cleaning ones needle, and daily rate that PWID exits population) to inform the parameters in our model. In addition, the prevalence of HCV among PWID in NYS is unknown and therefore cannot be used to validate the model. Although surveillance is currently conducted, risk factor data is not required for reporting purposes. Future studies conducted among PWID in NYS are needed and could be used to improve this study methodology. Fourth, the model does not specifically account for factors that may expedite the transition between disease stages such as alcohol use or HIV. Given the large HIV epidemic in NYS, this could impact the model. Further modeling incorporating these factors should be done, for example, by modeling HIV positive and HIV negative persons who consume alcohol to assess the difference. Lastly, the amount of equipment available was assumed to be a percentage of the number of needles dispensed or sold. Although sensitivity analyses explored varying this ratio, knowledge of a more precise amount would benefit the analysis.

Conclusion

The results of this analysis demonstrate the feasibility of using non-stationary quasi-Markov modeling to model the spread of HCV among PWID using free programming software. The R code in the appendix has been provided for other researchers to use and modify for their own purposes and could further impact this field of research as well as inform and promote additional education prevention and interventions for PWID.

Additional files

Competing interests

The authors declare that they have no competing interests.

Authors’ contributions

Acknowledgement

We would like to thank Colleen Flanigan, Director of the Viral Hepatitis Section at the AIDS Institute New York State Department of Health, for her contributions to the conception and interpretation of this study. We are grateful to Maxine Philips, Barbara Agatstein and the Division of Prevention at the New York State Department of Health for providing us pertinent information regarding syringe exchange programs and for their expertise in the field.

Publication history

Received: 27-Jun-2014 Final Revised: 27-Sep-2014 Accepted: 02-Oct-2014 Published: 07-Oct-2014

References


Citation: