



# HHS Public Access

Author manuscript

*Lancet Infect Dis.* Author manuscript; available in PMC 2017 June 01.

Published in final edited form as:

*Lancet Infect Dis.* 2016 June ; 16(6): 698–702. doi:10.1016/S1473-3099(16)00124-9.

## Phylogenetic estimation of the temporal spread of hepatitis C genotype 1a in North America: a retrospective phylogenetic study

Jeffrey B. Joy, PhD<sup>1,2,\*</sup>, Rosemary M. McCloskey, BSc<sup>1</sup>, Thuy Nguyen, BSc<sup>1</sup>, Richard H. Liang, PhD<sup>1</sup>, Yury Khudyakov, PhD<sup>3</sup>, Andrea Olmstead, PhD<sup>4</sup>, Mel Krajdén, MD, PhD<sup>4,5</sup>, John W. Ward, MD<sup>3</sup>, P. Richard Harrigan, PhD<sup>1,2</sup>, Julio S. G. Montaner, MD<sup>1,2</sup>, and Art F. Y. Poon, PhD<sup>1,2</sup>

<sup>1</sup>BC Centre for Excellence in HIV/AIDS, Vancouver, Canada

<sup>2</sup>Department of Medicine, University of British Columbia, Vancouver, Canada

<sup>3</sup>Division of Viral Hepatitis, Centers for Disease Control and Prevention, Atlanta, GA, United States

<sup>4</sup>BC Centre for Disease Control, Vancouver, Canada

<sup>5</sup>Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, Canada

### Abstract

**Background**—The timing of the initial spread of hepatitis C genotype 1a in North America is controversial. In particular how and when HCV reached extraordinary prevalence in specific demographic groups remains unclear. We sought to quantify, using all available HCV sequence data and phylodynamic methods the timing of the spread of HCV genotype 1a in North America.

**Methods**—We screened 45,316 publicly available HCV genotype 1a sequences for location and genotype then conducted phylogenetic analyses of available North American sequences from five HCV genes (E1, E2, NS2, NS4B, NS5B), with an emphasis on including as many sequences with early collection dates as possible. We inferred the historical population dynamics of this regional epidemic for all five HCV gene regions using Bayesian skyline plots.

**Findings**—Our analyses suggest that the bulk of the early spread of genotype 1a in North America occurred prior to 1965, and that the HCV epidemic has undergone relatively little expansion since 1965. Furthermore, our results indicate that the HCV epidemic has stabilized as of the early 1990's. These results were robust across all five HCV gene regions analyzed.

\*Corresponding author: BC Centre for Excellence in HIV/AIDS, 1081 Burrard Street, Vancouver, Canada, V6Z 1Y6, jjoy@cfenet.ubc.ca, +1 (604) 368-5569.

#### Author contributions

BJJ and AFYP designed the study. JBJ, TN, RMM, and AFYP acquired data. JBJ, RHL, AFYP analyzed and interpreted the data. JBJ and AFYP drafted the manuscript. All authors contributed to the critical revision of the manuscript and approved the final version.

#### Declaration of interests

None of the authors declare any conflict of interest with this study.

**Interpretation**—Our results reveal the early HCV epidemic dynamics in North America. Importantly, the expansion of genotype 1a prior to 1965 suggests that nosocomial or iatrogenic factors rather than past sporadic behavioral risk (i.e.: experimentation with injection drug use, unsafe tattooing, high risk sex, travel to high endemic areas) were key contributors to the HCV epidemic in North America. Our results may also reduce stigmatization around HCV screening and diagnosis, potentially increasing rates of HCV screening and treatment.

---

## Introduction

Hepatitis C virus (HCV) is a global threat to public health with an estimated 185 million infected persons worldwide, including an estimated 4.6 million infections of predominantly HCV genotype 1a in North America<sup>1</sup>. The North American epidemic is composed of an estimated 3.5 million infections in the United States, 300,000 infections in Canada, and 900,000 cases in Mexico<sup>1</sup>. Infected individuals are at elevated risk of liver cirrhosis, hepatocellular carcinoma, and liver failure, leading to earlier mortality and high cost to the health care system. About 75 percent of adults infected with HCV in North America are in the cohort composed of individuals born between 1945 and 1964 (the “baby boomers”)<sup>2</sup>. The mechanism by which HCV reached such high prevalence in this cohort remains unclear. Previous studies have implicated use of infected blood products prior to the advent of rigorous screening of the blood supply in the early 1990s for infectious agents and/or injection drug use which peaked in North America at the end of the 1960’s<sup>3,4</sup>. However, despite mixed evidence in the literature, the dominant view is that the baby boomer epidemic in North America is largely attributable to past sporadic behavioral risk (i.e., experimentation with injection drug use, unsafe tattooing, high risk sex, travel to high prevalence areas)<sup>2,5,6</sup> and transfusion with infected blood.

Many RNA viruses like HCV evolve rapidly and therefore it is possible to approximate the recent transmission history of an epidemic from the genetic divergence of infections sampled from the population<sup>7</sup>. This genetic diversity is structured by the shared ancestries of the infections, which can be reconstructed from these sequences as a tree-based model known as a phylogeny. Since most of the HCV genome evolves at ~0.001 substitutions/site/year<sup>8,9</sup>, it is possible to use the collection dates of different samples to rescale the phylogeny such that the branching points in the tree approximate the dates of transmission events in the epidemic<sup>10</sup>. Thus, the shape of the phylogeny contains valuable information about the epidemiological dynamics of the virus population<sup>11,12</sup>, which can be extracted by fitting epidemic or population genetic models to these data. Such techniques have previously provided key insights into ongoing epidemics including HIV, influenza A virus, dengue virus and Ebola virus<sup>4,7,13</sup>. Many of these studies have focused on the “effective number of infections”, which is a quantity expected to be proportional to prevalence during the exponential phase of an epidemic, but which should not in general be mistaken as equivalent to prevalence<sup>14</sup>.

In this study we use a distribution of phylogenetic trees of all published sequences covering five HCV genes to infer the past trajectories in the effective number of infections in the North American HCV genotype 1a epidemic. We explicitly focus on the genotype 1a epidemic because this is overwhelmingly the dominant HCV genotype circulating in North

America<sup>1</sup>. We examine the progression of the HCV epidemic in North America through time to test whether the timing of the spread of HCV in North America matches with the previously hypothesized periods of high incidence. If HCV epidemic dynamics in North America have primarily been driven by injection drug use and/or the use of infected blood products prior to the advent of effective screening of the blood supply for infectious agents, then we expect peak periods of the effective number of infections in North America to be highest, later in the progression of the HCV epidemic, at the end of the 1960's and early 1970's.

## Methods

### Data Collection

We retrieved all available hepatitis C virus sequence records from the National Center for Biotechnology Information (NCBI) Genbank nucleotide database (Appendix page 12). The search results, comprising 160,556 records, were reduced to 45,316 sequences by excluding any sequence lacking a source modifier field (attribute information associated with the sequence) or a collection date.

Specific gene regions were extracted from these data by pairwise alignment of each sequence against the respective sequence intervals from the H77 reference genome and extracting the overlapping region. Each gene-specific dataset was manually inspected and clipped to extract the broadest interval with the greatest overlap across sequences. We selected genomic regions for analyses based on previous whole genome studies of HCV genotype one which revealed E2, NS2, and NS5B as the most informative HCV gene regions for phylodynamic studies<sup>4</sup>. Additionally, we analyzed the E1 and NS4B regions<sup>4</sup>.

### Data Curation

To screen for genotype 1a sequences, genotype and gene specific reference sequences were collected from the Los Alamos National Laboratory HCV sequence repository. Sequences in each dataset were aligned using MAFFT v7.154b and manually inspected using HYPHY v2.22. Phylogenetic trees for each complete dataset were inferred in an approximate maximum likelihood modeling framework as implemented in FastTree2, then visualized and manually inspected using FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>) to isolate the clade comprising all sequences annotated as genotype 1a. This procedure identified 13 sequences whose GenBank records were incorrectly annotated with a different genotype. Subsequently, the resulting HCV gene-specific datasets comprised 917 sequences in total and covered 4,499 bases (E1:  $n=252$ , 576 bp; E2:  $n=196$ , 1,089 bp; NS2:  $n=182$ , 651 bp, NS4B:  $n=139$  NS4B, 783 bp; NS5B  $n=148$ , 1400 bp). Accession numbers for all public domain sequences used can be found in Appendix pages 5–10. Sampling times of included sequences ranged from 1977–2011 (Appendix Page 13). Approximate maximum likelihood phylogenetic trees for each gene region are displayed in Appendix Page 14.

### Effective Number of HCV Infections

We reconstructed the dynamics of the North American HCV epidemic over time using Bayesian skyline plots, a non-parametric smoothing method for approximating past

population dynamics<sup>15</sup>. The  $y$ -axis in a Bayesian skyline plot represents  $N_e$  multiplied by  $\tau$ , where  $N_e$  is the effective population size and  $\tau$  is the generation time. In this context the  $y$ -axis ( $N_e \times \tau$ ) is interpreted as the number of infected individuals who go on to infect additional individuals (effective number of infections<sup>16</sup>), and the  $x$ -axis represents time in years. A minimum of four replicate Markov chain Monte Carlo (MCMC) samples were run in BEAST v1.8.0 for each gene, for at least  $5 \times 10^8$  generations per run. To determine if choice of priors on both the model by which sequences accrue changes (substitution models) or the model by which the expected number of changes in a sequence relates to time (clock models) biased results, we performed two runs for each gene with four different substitution models (General Time Reversible (GTR) and Tamura-Nei 93 (TN93) with and without gamma distributed rate variation) and three different clock models (strict, random local, and relaxed). To determine the best fitting performed model comparison in the program Tracer v1.5 (<http://tree.bio.ed.ac.uk>) using Bayes Factors<sup>10</sup>. Optimal run conditions were determined to be a General Time Reversible substitution model with among site rate heterogeneity distributed according to a gamma distribution under a relaxed molecular clock. The convergence of each run was assessed through evaluation of parameter traces and effective sample size (ESS) values greater than 200 via the program Tracer v 1.5.

### Role of the Funding Source

The funders of this study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

## Results

### Sample Collection Dates

Due to the previous paucity of samples with early collection dates, previous studies focused on recovering HCV genotype 1a population dynamics utilized sequences with collection times heavily weighted towards the period between the year 2000 and 2010. For example, a landmark study used one sequence sampled before 1980 (H77), six before 1995, 44 between 2000 and 2004, and 35 falling between 2005 and 2007 (See Appendix Page 15). Due to recent improvements in sequencing technology and the technical ability to work with archival sample material more data is now available with earlier collection dates allowing a more even distribution of sampling times (See Appendix Pages 11, 13, and 15).

### Temporal Spread of HCV Genotype 1a in North America

Bayesian skyline plots reconstructing the timing of spread of the North American HCV genotype 1a epidemic for all analyzed HCV genomic regions are presented in Figure 1 and for each individual gene region in Appendix page 16. Taken together, the analyses of all gene regions suggest that the highest periods of epidemic expansion in North America occurred between 1940 and 1965. Importantly, our results show that the exponential growth phase of the epidemic had subsided by 1965, suggesting a plateau in the rate of spread between 1965 and 1989, then dropped during the early 1990's. Finally, Bayesian skyline plots suggest that there has been a slight increase in the HCV effective number of infections in the early 2000's. Analyses of each gene separately show substantial variation in estimates

of the timing of the early exponential growth ranging from approximately 1940, 1950, 1955, 1962, and 1965 for NS2, NS5B, E2, E1, and NS4B respectively. The most darkly shaded grey regions in Figure 1 represent time intervals where estimates of spread from each gene overlap in confidence. The time intervals across all genes where we had greatest confidence in the peak spread was located between 1948 and 1963. In sum, our phylogenetic analyses strongly suggest that the HCV genotype 1a epidemic in North America had already attained the height of its current distribution by 1960.

## Discussion

Our results indicate that the HCV genotype 1a epidemic underwent rapid expansion between 1940 and 1960 (Figure 1). There followed a period of relatively little growth in North American HCV epidemic, a subsequent decline during the early 1990's, and a small increase in the late 1990's early 2000's. Importantly, these analyses suggest the period of greatest increase of HCV genotype 1a in North America was substantially earlier than that previously suggested<sup>4</sup>.

Potential causes of a dramatic early and large-scale expansion (1940–1960) of the North American HCV genotype 1a epidemic are likely diverse, coinciding with the increase in medical procedures during and immediately following World War II<sup>4,17,18</sup> when injection and blood transfusion technology was still in its infancy<sup>18,19</sup>, and the expansion of recreational injection drug use in North America and associated needle sharing between 1920 and the late 1960s<sup>3,5</sup>. Many medical procedures have been linked to the spread of HCV including HCV contaminated multi-dose vials<sup>20–22</sup>, finger stick devices<sup>23</sup>, and surgical procedures<sup>24</sup>. The spread of HCV in North America through iatrogenic sources between 1940–1960 is consistent with studies revealing the iatrogenic spread of HCV in other regions of the developed world in the same time frame; for example the post World War II spread of HCV in Russia<sup>25</sup>, France<sup>26</sup>, and Japan<sup>18,25</sup>. Injection drug use increased in frequency in North America in the late 1960s suggesting that if this was the main driver of the epidemic the period of peak expansion of the HCV epidemic would have coincided with the peak of injection drug use rather than much earlier as shown by our analyses.

The observed plateau in HCV spread between 1960 and 1990 is consistent with the hypothesis that changes in injection technology were a driving factor. Prior to 1950, injection technology was characterized by machine made glass and metal syringes which were typically manually sterilized and reused due to their high cost<sup>19</sup>. A transition period occurred between 1950–1960, when such glass and metal syringes were phased out and replaced by disposable plastic single-use syringes<sup>19</sup>. Since 1960 medical syringe reuse has been markedly reduced in North America<sup>19</sup>. The HCV epidemic spread during this plateau period may have been maintained by a combination of injection drug use<sup>3,4</sup> and transfusion of infected blood products prior to the commencement of rigorous screening of the blood supply in 1992<sup>27</sup>. Before 1992, the incorporation of plasma infected with HCV into the blood supply commonly resulted in nosocomial HCV transmission during medical procedures involving the transfusion of blood or organ transplantation<sup>27</sup>. However, HCV transmission via infected blood products combined with transmission through injection drug use likely occurred at a rate sufficient to maintain the effective population size of HCV<sup>3,5,6</sup>.

Further support for this hypothesis is seen in the observed decline in spread in the early 1990's coincident with the advent of rigorous screening of the blood supply<sup>27</sup> and with the widespread penetration of needle exchange programs into injection drug user communities<sup>28</sup>. The slight upturn in effective population size observed just after the year 2000 is consistent with epidemiological evidence of both increases in HCV infection amongst young injection drug users who reside in non-urban areas<sup>29</sup>, and a significant increase in HCV infections amongst men who have sex with men<sup>30</sup>.

A similar analysis by Magiorkinis *et al.*<sup>4</sup>, focused on the NS5B gene, estimated the main expansion of the global HCV genotype 1a epidemic to be between 1960 and 1980 peaking in 1975. Our reanalysis of the Magiorkinis *et al.* data, using similar methods but restricted to samples collected in the United States, was consistent with their trend. The observed discordance between the timing of the spread of HCV genotype 1a in North America in our study and that by Magiorkinis *et al.* most likely reflects the greater number of early sequences included in our analyses (See Appendix pages 11, 13, and 15) and the fact that this study is restricted to North American sequences whereas Magiorkinis *et al.* included sequences from other regions of the world. Variation in the rate at which changes in nucleotide sequences (substitutions) accrue with respect to time is observed between lineages or between time points. This observed variation in the rate at which substitutions accrue relative to time may occur because of fluctuations in population size, host-specific environment effects, the intensity of natural selection, changes in protein function, or changes in generation times. Thus, if most of the sampling times are clustered in later time windows, this can bias estimates in the direction of the clump of sampling times. A more even distribution of sampling times will provide more robust estimates that are less influenced by such variation in the rate at which molecular changes accumulate in genetic sequences.

Based on our results, the oldest members of the demographic cohort displaying the highest burden of HCV (the “baby boomers”) were roughly five years of age by 1950, around the peak of the spread of genotype 1a in North America. Thus, it is unlikely that past sporadic behavioral risk (experimentation with injection drug use, unsafe tattooing, high risk sex, travel to high endemic areas) was the dominant route of HCV transmission in this demographic. Though HCV infection in this cohort through injection drug use certainly occurred thereafter, and remains a dominant source of new infections<sup>5,6</sup>.

Limitations of this analysis include the fact that phylogenetic methods of dating ancestors from genetic sequences are sensitive to invalid model assumptions<sup>13</sup> making it problematic to confidently pinpoint a difference of 10–20 years deep in the past. The inclusion of a greater number of older HCV sequences obtained from archived patient samples would be of great value in refining estimates of the timing of the rate of spread while simultaneously improving confidence intervals around the estimates. It is not unusual for molecular clock estimates to have broad confidence intervals when reconstructing events deep in the tree; furthermore, our gene specific datasets do not overlap completely. Thus, we are integrating sequences derived from different genomes that have been subjected to different environments. We therefore, should not expect the confidence limits on these different gene specific datasets to be highly concordant in the remote past.



## Conclusions

In summary, our data indicate the rapid and large-scale expansion of HCV transmission in North America was coincident with increases in medical procedures that began after World War II (nosocomial or iatrogenic causes), and not only the rise in injection drug use, which peaked much later in North America, in the late 1960's<sup>3</sup>. Thus, the prevailing view that in North America the HCV epidemic is predominately attributable to past sporadic behavioral risk (i.e.: experimentation with injection drug use, unsafe tattooing, high risk sex, travel to high endemic areas) is not supported by our data. Belief that the dominant route to HCV infection in North America has been behavioural risk may introduce a bias amongst providers and patients against HCV screening. The results presented here may help to reduce stigmatization around HCV screening and diagnosis, potentially increasing the numbers of providers offering screening, patients accepting, and if positive, presenting for care and treatment. Importantly, because of vast improvements in medical technology through time, particularly in syringe manufacture and guidelines for their use, nosocomial factors which contributed to the early spread of HCV in North America<sup>4</sup> no longer play a major role in the epidemic.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

### Funding

The Canadian Institutes of Health Research, Michael Smith Foundation for Health Research, and BC Centre for Excellence in HIV/AIDS.

## Literature Cited

1. Messina JP, Humphreys I, Flaxman A, et al. Global distribution and prevalence of hepatitis C virus genotypes. *Hepatology*. 2015; 61(1):77–87. [PubMed: 25069599]
2. Cohen J. Calling all baby boomers: get your hepatitis C test. *Science*. 2012; 337:903. [PubMed: 22923556]
3. Courtwright, DT., Courtwright, DT. *Dark paradise: A history of opiate addiction in America*. Harvard University Press; 2009.
4. Magiorkinis G, Magiorkinis E, Paraskevis D, et al. The Global Spread of Hepatitis C Virus 1a and 1b: A Phylodynamic and Phylogeographic Analysis. *PLoS Med*. 2009; 6(12):e1000198. [PubMed: 20041120]
5. Williams IT, Bell BP, Kuhnert W, Alter MJ. Incidence and transmission patterns of acute hepatitis C in the United States, 1982–2006. *Archives of internal medicine*. 2011; 171(3):242–248. [PubMed: 21325115]
6. Alter MJ. Epidemiology of hepatitis C. *Hepatology*. 1997; 26(S3):62S–65S. [PubMed: 9305666]
7. Poon AFY, Joy JB, Woods CK, et al. The Impact of Clinical, Demographic and Risk Factors on Rates of HIV Transmission: A Population-based Phylogenetic Analysis in British Columbia, Canada. *Journal of Infectious Diseases*. 2014:jiu560.
8. Gray RR, Parker J, Lemey P, Salemi M, Katzourakis A, Pybus OG. The mode and tempo of hepatitis C virus evolution within and among hosts. *BMC evolutionary biology*. 2011; 11(1):131. [PubMed: 21595904]

9. Kimura M. Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. *Nature*. 1977; 267:275–276. [PubMed: 865622]
10. Rodrigo AG, Felsenstein J. Coalescent approaches to HIV population genetics. The evolution of HIV. 1999:233–272.
11. Stadler T, Bonhoeffer S. Uncovering epidemiological dynamics in heterogeneous host populations using phylogenetic methods. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2013; 368(1614):20120198.
12. Holmes EC, Nee S, Rambaut A, Garnett GP, Harvey PH. Revealing the History of Infectious Disease Epidemics through Phylogenetic Trees. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*. 1995; 349(1327):33–40. [PubMed: 8748017]
13. Worobey M, Rambaut A, Han G-Z. A synchronized global sweep of the internal genes of modern avian influenza virus. *Nature*. 2014; 508:254–257. [PubMed: 24531761]
14. Frost SDW, Volz EM. Viral phylodynamics and the search for an ‘effective number of infections’. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2010; 365(1548):1879–1890.
15. Drummond AJ, Rambaut A, Shapiro B, Pybus OG. Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular biology and evolution*. 2005; 22(5): 1185–1192. [PubMed: 15703244]
16. Frost SDW, Volz EM. Modelling tree shape and structure in viral phylodynamics. *Philosophical Transactions of the Royal Society Series B*. 2013; 368:1614.
17. McNaughton AL, Cameron ID, Wignall-Fleming EB, et al. Spatiotemporal Reconstruction of the Introduction of Hepatitis C Virus into Scotland and Its Subsequent Regional Transmission. *Journal of Virology*. 2015; 89(22):11223–11232. [PubMed: 26311892]
18. Prati D. Transmission of hepatitis C virus by blood transfusions and other medical procedures: a global review. *Journal of hepatology*. 2006; 45(4):607–616. [PubMed: 16901579]
19. Drucker E, Alcabes PG, Marx PA. The injection century: massive unsterile injections and the emergence of human pathogens. *The Lancet*. 2001; 358:1989–1992.
20. Krause G, Trepka MJ, Whisenhunt RS, et al. Nosocomial transmission of hepatitis C virus associated with the use of multidose saline vials. *Infection Control*. 2003; 24(02):122–127.
21. Widell A, Christensson B, Wiebe T, et al. Epidemiologic and molecular investigation of outbreaks of hepatitis C virus infection on a pediatric oncology service. *Annals of internal medicine*. 1999; 130(2):130–134. [PubMed: 10068359]
22. Tallis GF, Ryan GM, Lambert SB, et al. Evidence of patient - to - patient transmission of hepatitis C virus through contaminated intravenous anaesthetic ampoules. *Journal of viral hepatitis*. 2003; 10(3):234–239. [PubMed: 12753344]
23. Petit JM, Bour JB, Aho LS, Castaneda A, Vaillant G, Brun JM. HCV and diabetes mellitus: influence of nosocomial transmission with the use of a finger stick device. *The American journal of gastroenterology*. 1999; 94(6):1709–1710. [PubMed: 10364057]
24. Ross RS, Viazov S, Roggendorf M. Phylogenetic analysis indicates transmission of hepatitis C virus from an infected orthopedic surgeon to a patient. *Journal of medical virology*. 2002; 66(4): 461–467. [PubMed: 11857522]
25. Tanaka Y, Kurbanov F, Mano S, et al. Molecular tracing of the global hepatitis C virus epidemic predicts regional patterns of hepatocellular carcinoma mortality. *Gastroenterology*. 2006; 130(3): 703–714. [PubMed: 16530512]
26. Henquell C, Guglielmini J, Verbeeck J, et al. Evolutionary history of hepatitis C virus genotype 5a in France, a multicenter ANRS study. *Infection, Genetics and Evolution*. 2011; 11(2):496–503.
27. Rich JD, Taylor LE. The beginning of a new era in understanding hepatitis C virus prevention. *Journal of Infectious Diseases*. 2010; 202(7):981–983. [PubMed: 20726769]
28. Des Jarlais DC, Perlis T, Arasteh K, et al. Reductions in hepatitis C virus and HIV infections among injecting drug users in New York City, 1990–2001. *Aids*. 2005; 19:S20–S25. [PubMed: 16251819]
29. Valdiserri R, Khalsa J, Dan C, et al. Confronting the emerging epidemic of HCV infection among young injection drug users. *American journal of public health*. 2014; 104(5):816–821. [PubMed: 24625174]



30. Hagan H, Jordan AE, Neurer J, Cleland CM. Incidence of sexually transmitted hepatitis C virus infection in HIV-positive men who have sex with men. *AIDS*. 2015; 29(17):2335–2345. [PubMed: 26258525]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Panel: Research in context****Evidence before this study**

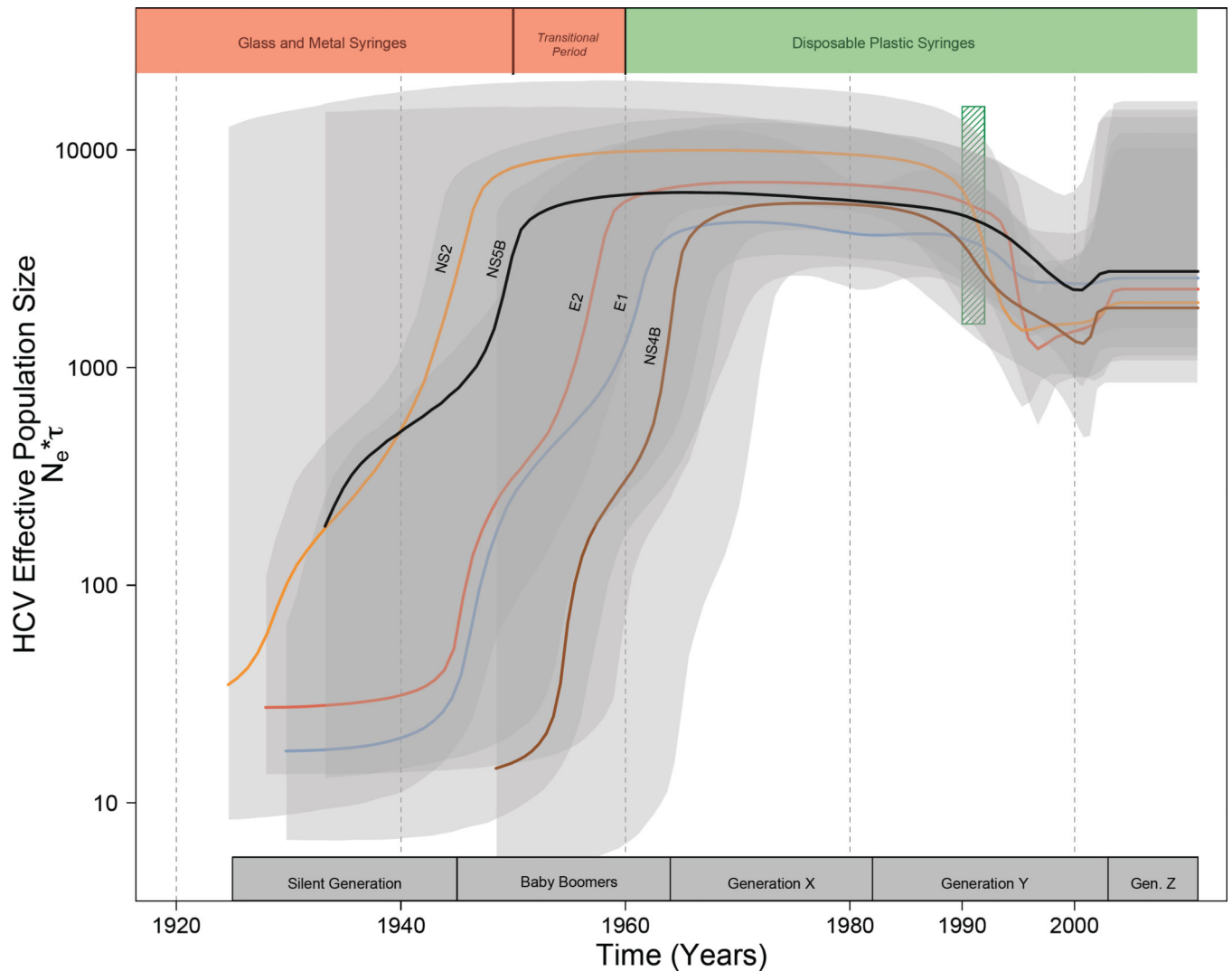
We searched Scopus, PubMed and Google Scholar for articles published up to December 2, 2015, in any language, that addressed the spread of hepatitis C in North America and globally. The vast majority of adults infected with HCV in North America, and for some other developed countries, are in the cohort composed of individuals born between 1945 and 1964 (the “baby boomers”). How HCV reached such high prevalence in this cohort remains unclear. Previous studies implicated both use of infected blood products prior to screening of the blood supply in the early 1990s and/or injection drug use, which peaked in North America at the end of the 1960’s. However, despite mixed evidence in the literature, the HCV epidemic in North America is thought to be largely attributable to past sporadic behavioral risk (*i.e.*, experimentation with injection drug use, unsafe tattooing, high risk sex, travel to high prevalence areas), along with limited transmission through transfusion with infected blood products.

**Added value of this study**

Our results, based on careful phylogenetic analyses within a large sample of North American HCV sequences, revise previous estimates of the exponential phase of growth of the North American HCV genotype 1a epidemic back in time by more than 15 years to approximately 1950. These results dispute the idea that the HCV epidemic amongst Baby Boomers and other demographic groups in North America is primarily due to injection drug use, unsafe tattooing, high-risk sex, and travel to high endemic areas during early youth. It instead suggests that the epidemic is linked to both the increase in medical procedures that occurred during and following World War II and the use of reusable glass and metal syringes in the 50’s and 60’s and subsequently behavioural risk.

**Implication of all the available evidence**

The spread of HCV genotype 1a in North America and dominance in one particular demographic cohort (“the baby boomers”) is a result of early expansion due to iatrogenic and behavioural risk, followed by transmission due to intravenous drug use in later years as medical technology and blood supply screening prevented iatrogenic transmission.



**Figure 1.** Coloured labeled lines for each gene region represent the median estimated HCV effective number of infections (effective number of infections  $\times$  generation time;  $N_e \tau$ ) through time estimated under a relaxed-clock model. The 95% highest probability density (HPD) for each gene region is shaded in grey. Areas of darker grey represent areas with the most HPD overlap across gene regions. Birth intervals for each generation are denoted along the  $x$ -axis. The vertical hatched green box corresponds to the advent of both rigorous screening of the blood supply for infectious agents in 1992<sup>27</sup> and the widespread availability of needle exchange programs in North America<sup>28</sup>. Coloured polygons along the upper portion of the plot denote the history of syringe use in North America<sup>19</sup>. Note the  $y$ -axis is on a log-scale units unaffected.