

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39

## Molecular and spatial epidemiology of HCV among people who inject drugs in Boston, Massachusetts

Thomas J. Stopka, PhD, MHS,<sup>1</sup> Omar Yaghi, MD, MPH,<sup>1</sup> Min Li, PhD,<sup>2</sup> Elijah Paintsil, MD,<sup>2,3</sup> Kenneth Chui, PhD, MPH,<sup>1</sup> David Landy, MSPH,<sup>1</sup> and Robert Heimer, PhD<sup>2,4</sup>

### Affiliations

<sup>1</sup>Department of Public Health and Community Medicine, Tufts University School of Medicine

<sup>2</sup>Yale University School of Public Health

<sup>3</sup>Yale New Haven Hospital

<sup>4</sup>Center for Interdisciplinary Research on AIDS at Yale

### Corresponding Author:

Robert Heimer, PhD

Professor of Epidemiology

Department of Epidemiology of Microbial Diseases

Yale University School of Public Health

60 College St.

PO Box 208034

New Haven, CT 06520-8034

Phone: (203) 785-6732

Email: [robert.heimer@yale.edu](mailto:robert.heimer@yale.edu)

Running Title: HCV Molecular Spatial Epidemiology: PWID

**Abstract Word Count:** 100

**Manuscript Word Count:** 2,020

**Tables/Figures:** 2

**References:** 17

40

41

## **Abstract**

42

43 Integration of genetic, social network, and spatial data has the potential to improve  
44 understanding of transmission dynamics in established HCV epidemics. Sequence data  
45 were analyzed from 63 viremic people who inject drugs recruited in the Boston area through  
46 chain referral or time-location sampling. HCV subtype 1a was most prevalent (57.1%),  
47 followed by subtype 3a (33.9%). The phylogenetic distances between sequences were no  
48 shorter comparing individuals within versus across networks, nor by location or time of first  
49 injection. Social and spatial networks, while interesting, may be too ephemeral to inform  
50 transmission dynamics when the date and location of infection are indeterminate.

51 **Keywords:** HCV, people who inject drugs, phylogenetic analyses, spatial analyses, social  
52 networks.

53

54

## 55 INTRODUCTION

56 Although now curable, hepatitis C virus (HCV) infections continue to be problematic.  
57 In the United States, HCV mortality has increased in recent years, while mortality for 61  
58 other federally reportable infectious diseases, including HIV, has decreased.<sup>1</sup> HCV infections  
59 have increased by more than 70% among 15-30 year-olds during the past decade, and new  
60 infections are largely attributed to injection drug use (IDU).<sup>1,2</sup> In Massachusetts, over 2,000  
61 new cases attributed to IDU have been reported in those <30 years of age in each of the last  
62 seven years.<sup>3</sup> More information about transmission networks might improve prevention  
63 efforts.

64 The application of phylogenetic and spatial analysis of HCV has identified HCV  
65 subtypes among people who inject drugs (PWID) and provided biological evidence of  
66 disease clusters and transmission patterns in outbreaks.<sup>4-6</sup> We aimed to find evidence of  
67 disease clusters or transmission patterns among PWID outside of known outbreaks. As part  
68 of a study of PWID in the Boston area, we identified a subsample with active HCV infection  
69 and sequenced part of their viral genome. We then integrated behavioral, social network,  
70 spatial, and HCV phylogenetic data to explore transmission dynamics in established  
71 epidemics. We were especially interested to see if we could find evidence of disease  
72 clusters or transmission patterns that could identify potential targets, methods, or venues for  
73 HCV prevention and treatment interventions.

## 74 METHODS

75 We employed a cross-sectional design focused on PWID in the Boston area. PWID  
76 were eligible if they met the following criteria: (1) age 15-45 years; (2) English or Spanish  
77 speaking; (3) evidence of track marks; (4) living with, at risk for, or infected with and  
78 subsequently cleared HCV; and (5) ability to provide written informed consent. The study  
79 protocol was approved by the [Blinded] Institutional Review Board and the [Blinded] Human

80 Investigation Committee. Consenting individuals completed a survey and were administered  
81 the Oraquick® HCV Rapid Antibody test (OraSure Technologies).

82           Between February and October 2016, recruitment relied on time-location sampling  
83 and chain referral in Boston and Cambridge. For the ***time-location sampling*** approach,  
84 trained outreach staff recruited participants through partnership with harm reduction  
85 programs, shelters, and social service agencies. Staff members at community-based  
86 agencies posted study flyers and identified PWID whom they knew were likely to have  
87 acquired HCV and would be interested in participating in the study. For the ***chain-referral***  
88 ***approach***, enrolled participants (i.e., “seeds”) were asked to refer peers within their  
89 networks to the study. In both approaches, research staff described the study to interested  
90 recruits, confirmed participant eligibility, provided written materials, and obtained written  
91 informed consent. Analysis of recruitment chains was depicted using the StatNet and  
92 UserNetR packages in R.

93           Participants were tested for HCV antibodies and then completed the survey on a  
94 tablet, laptop, or desktop computer in our research offices or on tablets or laptops in private  
95 spaces provided by community-based partners. Surveys covering demographic, health, and  
96 socio-behavioral topics were conducted in English or Spanish and lasted approximately 60  
97 minutes. Participants were informed of HCV test results after completing the survey, and  
98 those who tested positive were invited to a separate appointment and were consented to  
99 have a blood sample taken. Sample collection continued until 102 individuals were enrolled  
100 in the sub-study of HCV-positive participants.

101           Blood specimens were centrifuged, and the serum was added to RNALater® to  
102 preserve the integrity of viral RNA (RNALater; Ambion Inc., Austin, TX). Samples were  
103 stored at -80°C. After thawing, viral RNA was purified and a 360-base region in the core  
104 gene was amplified and sequenced to identify genotypes and produce phylogenetic trees.  
105 We optimally aligned the nucleotide sequences using the CLUSTAL W program and  
106 constructed a phylogenetic tree by the Neighbor-Joining method based on Kimura’s two-

107 parameter distances (Mega software, version 7.0). We assessed the reliability of the tree  
108 topologies by bootstrapping with 100 replicates, with a bootstrap agreement of  $\geq 70\%$   
109 required to define a phylogenetic cluster. Evolutionary distances were computed using the  
110 Maximum Composite Likelihood method.<sup>7</sup>

111 We compared sequence similarity by recruitment pattern, location of injection  
112 initiation, and number of years injecting drugs. We conducted Fisher's exact tests to assess  
113 whether differences in HCV subtypes by recruitment approach were statistically significant.  
114 We subsequently ran simple logistic regression analyses to determine whether place of  
115 residence at the time of first injection or number of years injecting was significantly  
116 associated with HCV subtypes. Place of residence was divided into three categories:  
117 Boston, suburbs of Boston (i.e., outside of the city limits and on a Massachusetts Bay Transit  
118 Authority (MBTA) line), or beyond the suburbs (not reachable by MBTA line). The number of  
119 years injecting was also divided into three categories: 0-5, 6-10, >10 years.

## 120 **RESULTS**

121 Among the 102 participants who tested positive for HCV antibodies and provided a  
122 subsequent blood sample, 66 were actively infected and sequence data were obtained from  
123 63 of the 66. Out of these 63, 60 completed the survey, and one additional participant  
124 completed the survey without providing sequence data. Thus, the sample sizes are 63 for the  
125 sequence data and 61 for the survey. The mean age of the 61 respondents was  $32 \pm 6$  years,  
126 and males comprised 75% of the sample. Forty (66%) described themselves as White non-  
127 Latino, nine as Latino, eight as mixed-race or other, and four as African American. Most were  
128 single, with 42 (69%) living alone and 9 (15%) living with a partner. About three-quarters  
129 (n=46) reported their sexual orientation as heterosexual, nine as bisexual, and six as gay or  
130 lesbian. Three-quarters of the sample reported being homeless and the median monthly  
131 income was \$500-1000. Ninety percent (n=53) reported having been incarcerated. The mean  
132 duration of IDU was 12 years (SD=8) (Table 1).

### 133 ***HCV subtype by recruitment pattern***

134 Figure 1A shows the participant network. Thirty-one of the 63 participants were  
135 identified through chain referral, indicated by being connected with double-headed arrows,  
136 and 32 were identified through time-location sampling, indicated by nodes without  
137 connection. Genotype 1a was higher for time-location participants (56.3%; 18/32) than for  
138 chain referred participants (41.9%; 13/31; p-value: 0.211); genotype 3a was higher for chain  
139 referred participants (45.2%; 14/31) than for time-location participants (22.6%; 7/32; p-value:  
140 0.109). While there was no exclusive clustering of subtypes within the chain referral clusters,  
141 three clusters with  $n > 3$  only consisted of two subtypes. In addition, in the largest clusters  
142 ( $n=13$ ), seven participants with subtype 3a were connected.

### 143 ***HCV subtype by location of first injection***

144 Among PWID with Subtype 1a, 32.4% (11/34) first injected in Boston, 20.6% (7/34)  
145 first injected in Boston suburbs, and 47.1% (16/34) first injected outside of the Boston  
146 suburbs. Among PWID with subtype non-1a, 39.1% (9/23) first injected in Boston, 30.4%  
147 (7/23) first injected in Boston suburbs, and 30.4% (7/23) first injected outside of the suburbs.  
148 Location of the first injection was not associated with HCV subtype, when comparing  
149 subtypes from outside of the Greater Boston Area (odds ratio [OR]: 1.87; 95% confidence  
150 interval [CI]: 0.54, 6.53) or in Boston suburbs (OR: 0.82; 95% CI: 0.21, 3.26) to subtypes  
151 from Boston.

### 152 ***HCV subtype by number of years injecting***

153 Among PWID with Subtype 1a, 21.9% (7/32) had been injecting for 0-5 years, 31.3%  
154 (10/32) had injected for 6-10 years, and 46.9% (15/32) had been injecting for >10 years.  
155 Among PWID with Subtype Non-1a, 13.7% (3/22) had been injecting for 0-5 years, 13.7%  
156 (3/22) had injected for 6-10 years, and 72.7% (16/22) had been injecting for >10 years. We  
157 divided the sample into those infected with genotype 1a versus not 1a, and there was no  
158 difference between groups by years of injection experience (OR: 0.97, 95% CI: 0.90, 1.04).

159 ***HCV sequencing results***

160 Sequencing results indicated that the distances between sequences in phylogenetic  
161 analyses were no shorter comparing sequence data within versus across networks (Figure  
162 1B). Four of the oldest injectors in the sample (individuals born in 1977 or earlier) were  
163 infected with HCV genotypes 1b and 2b. Only one other individual was infected with either of  
164 these two subtypes (Figure 1B).

165 **DISCUSSION**

166 Integrating social network, spatial, and molecular data from established epidemics  
167 poses unique challenges. These stem from the lack of defined dates of infection, the  
168 durability of current social network connections, the multilevel nature of risk environments,<sup>8,9</sup>  
169 and missing data about the full network structure that is obtained when chain referral  
170 relationships are the sole data source on social connections. For this reason, we sought to  
171 determine if HCV was more closely related (1) when specimens were obtained from within  
172 recruitment chains than those obtained by the less directed recruitment achieved by time-  
173 location sampling, and (2) when participants were categorized by time and location that their  
174 first exposure to HCV had likely have occurred. For the latter analysis, we made the working  
175 hypothesis that HCV infection is an early consequence of initiating (IDU). This hypothesis is  
176 consistent with observations of high HCV incidence in the first year of injection.<sup>10,11</sup>

177 Our findings that HCV subtype 1a was the predominant subtype among PWID in the  
178 Boston Area (57.1%), followed by subtype 3a (33.9%), differs from studies elsewhere that  
179 focused on HCV among PWID. A study from New York, conducted twenty years earlier than  
180 ours, found that the predominant genotypes were 1a and 1b, with 3a comprising only 5% of  
181 sequences genotyped.<sup>12</sup> A study from Baltimore, concurrent with ours, found no evidence of  
182 genotype 3a.<sup>13</sup> Our study population was substantially younger than those in the other two  
183 studies, consistent with our aim to recruit a younger sample, but differences in genotype  
184 were not associated with age. However, there were two subtypes (1b and 2b) that were

185 restricted to older individuals. It appears, however, that these subtypes have failed to spread  
186 among younger members of our study population. Analysis of HCV genotypes from PWID in  
187 Vancouver revealed a pattern similar to ours, with nearly half their sample infected with  
188 genotype 1a and one-third infected with genotype 3a, but the relevance of this finding is  
189 limited by the geographic distance between coasts.<sup>6</sup>

190 HCV phylogenetic clustering has been found in those recently initiating drug injection.  
191 Clustering was quite common in the large HIV-HCV outbreak in Scott County, Indiana. But  
192 even there, multiple introductions of HCV were identified. Genotypes 1a and 3b accounted  
193 for 72% and 20% of the patient samples, respectively, but additional genotypes had entered  
194 the mix, and some people were infected with multiple genotypes.<sup>14</sup> One study from  
195 Melbourne found within-network correlation with HCV genotype, but only in individuals  
196 infected with genotype 3 and reporting recent drug injection initiation.<sup>15</sup> In our study, we did  
197 not find any significant associations between HCV subtypes by recruitment types, number of  
198 years injecting, nor location of first injection.

199 Several limitations should be considered in interpreting our findings. First, our sample  
200 is small and needs to be expanded. Second, while we aimed to use respondent driven  
201 sampling (RDS), we ultimately used a modified RDS approach combined with a time-location  
202 sampling approach to recruit our sample during a relatively tight timeframe based on funding  
203 requirements. Third, we used self-reported measures from our survey. Some questions we  
204 asked referred to events from the distant past (e.g., place where first injected, date of first  
205 injection), which could have introduced recall bias. Fourth, our results are not generalizable  
206 to the entire Boston Area, nor to other cities. Finally, although we sought to recruit a  
207 relatively young sample of PWID, many people had long injecting experiences, so their  
208 social networks at the time they were infected with HCV may have been different from their  
209 social networks at the time of the survey. In lieu of trying to design studies that collect  
210 detailed individual network and phylogenetics to reconstruct patterns of disease  
211 transmission, it may be more reasonable to collect information sufficient to populate random



212 mixing and other forms of stochastic models.<sup>16</sup> Sequencing of infectious disease  
213 transmission among PWID may be most effective in the case of disease outbreaks.<sup>17</sup>

## 214 **Conclusion**

215 Social networks, while interesting, are too ephemeral to inform transmission  
216 dynamics if the date and location of infection are indeterminate. Expanding research efforts  
217 to obtain extensive social network data from PWID populations in established epidemics  
218 may not be worthwhile or cost-effective given the expense of obtaining full social network  
219 data. Alternatives, such as quasi network data obtained through RDS and random mixing  
220 models, may be sufficient when modeling transmission networks.

221

## 222 **Acknowledgments**

223 The authors wish to acknowledge the assistance of the following agencies and individuals  
224 who contributed to recruitment and data collection efforts through participant referrals or  
225 provision of space for data collection activities: AHOPE, AIDS Action Committee, South  
226 Boston Behavioral Health Center, Dr. Laura Grubb, Tufts Medical Center, and St. Anthony's  
227 Shelter. The authors also wish to thank Dr. Nadia Abdala for guidance on HCV specimen  
228 sequencing and preparation for phylogenetic analyses.

## 229 **Funding Support**

230 This research was funded by a 2016 grant from the Tufts Institute for Innovation (PI:  
231 Stopka). Additional support was provided by the Providence-Boston Center for AIDS  
232 Research (PI: Stopka), through Grant Number P30AI042853 from the National Institute of  
233 Allergy and Infectious Diseases. Support for Dr. Heimer was provided through Grant Number  
234 R01DA030420 from the National Institute on Drug Abuse. The content is solely the  
235 responsibility of the authors and does not necessarily represent the official views of the  
236 National Institute of Allergy and Infectious Diseases or the National Institute of Health.

237

238 **Conflicts**

239 The authors report no conflicts of interest.

240

241 **Contributions**

242 TS conceptualized the study, oversaw all study activities, wrote the initial draft of the  
243 manuscript, and obtained the funding for the study. OY assisted with field operations, data  
244 collection, data management, and review of the final manuscript. ML conducted phylogenetic  
245 analyses, development of data visualizations, and interpretation of findings. EP oversaw all  
246 laboratory analyses and interpretation of phylogenetic analyses. KC conducted statistical  
247 analyses and assisted with the development of all figures. DL coordinated field activities,  
248 including recruitment, data collection, and data management. RH contributed to study  
249 conceptualization, the analytical plan, phylogenetic analyses, and interpretation of findings.  
250 All authors contributed to, reviewed, and approved the final version of the manuscript.

251

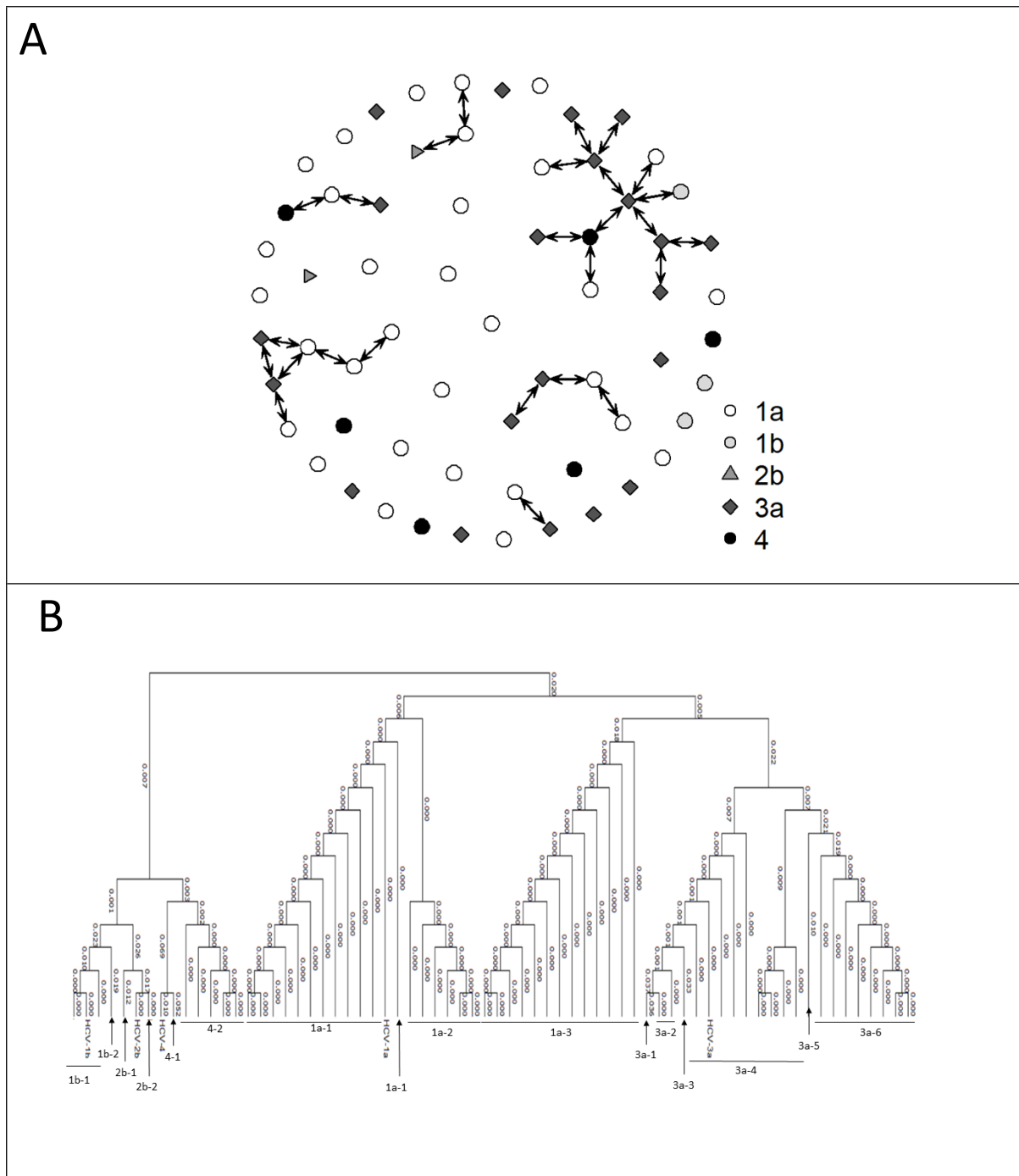
252  
253

**Table.** Descriptive statistics: People who inject drugs in the Greater Boston Area of Massachusetts, 2016 (n=61).

	<b>Total</b>
	N=61
Age (year)	32 (6)
Gender	
Male	46 (75%)
Female	15 (25%)
Race/ethnicity	
African American	4 (7%)
American Indian/Alaskan Native	1 (2%)
Hispanic, Latino, or Latina	9 (15%)
White, non-Hispanic	40 (66%)
Mixed races & others	7 (11%)
Marital status	
Single/living alone	42 (69%)
Single, living w/ partner	9 (15%)
Married/ in domestic partnership	2 (3%)
Divorced	3 (5%)
Separated	3 (5%)
Widowed	2 (3%)
Sexual orientation	
Heterosexual/straight	46 (75%)
Gay/lesbian	6 (10%)
Bisexual	9 (15%)
Education level	
Grade 6-8	3 (5%)
Grade 9-12, not HS grad	14 (23%)
HS grad/GED	32 (52%)
Some colleage, no degree	9 (15%)
2-4 yrs college degree	1 (2%)
Graduate/professional degree	2 (3%)
Homeless	
No	15 (25%)
Yes	46 (75%)
Main residence, last 30 days	
Own apartment, house, or room	7 (11%)
Home of parents, relatives, or friends	10 (16%)
Halfway house, group home, or foster home	6 (10%)
Hotel or motel	2 (3%)
Shelter	18 (30%)
Abandoned building	2 (3%)
Public park	4 (7%)
Street, wooded area	7 (11%)
Other	5 (8%)
Income, last 30 days	
None	7 (11%)
< \$500	17 (28%)
\$500 to < \$1000	22 (36%)
\$1000 to < \$2000	8 (13%)
\$2000 to < \$4000	4 (7%)
\$4000 to < \$6000	1 (2%)
>= \$6000	2 (3%)
Ever in jail, prison, JDC	
No	6 (10%)
Yes	55 (90%)
Currentlly on parole	
No	53 (87%)
Yes	8 (13%)

254

255



256

257 **Figure 1.** A) Network diagram showing the referral network. The different node symbols  
258 represent the five major hepatitis C virus (HCV) subtypes. The double-headed arrows  
259 indicate connection by referral (n=31). B) Phylogenetic tree for HCV positive people who  
260 inject drugs in Boston, Massachusetts, 2016 (n=63).

261

## 262 References

- 263 1. Ly KN, Hughes EM, Jiles RB, Holmberg SD. Rising Mortality Associated With  
264 Hepatitis C Virus in the United States, 2003-2013. *Clin Infect Dis*. May 15 2016;62(10):1287-  
265 1288. doi:10.1093/cid/ciw111
- 266 2. Suryaprasad AG, White JZ, Xu F, et al. Emerging epidemic of hepatitis C virus  
267 infections among young nonurban persons who inject drugs in the United States, 2006-2012.  
268 *Clin Infect Dis*. Nov 2014;59(10):1411-9. doi:10.1093/cid/ciu643
- 269 3. Onofrey S, Aneja J, Haney GA, et al. Underascertainment of acute hepatitis C virus  
270 infections in the U.S. surveillance system: a case series and chart review. *Ann Intern Med*.  
271 Aug 18 2015;163(4):254-61. doi:10.7326/m14-2939
- 272 4. Mathei C, Van Dooren S, Lemey P, Van Damme P, Buntinx F, Vandamme AM. The  
273 epidemic history of hepatitis C among injecting drug users in Flanders, Belgium. *J Viral*  
274 *Hepat*. Jun 2008;15(6):399-408. doi:10.1111/j.1365-2893.2007.00950.x
- 275 5. Caraballo Cortes K, Rosinska M, Janiak M, et al. Next-generation sequencing  
276 analysis of a cluster of hepatitis C virus infections in a haematology and oncology center.  
277 *PLoS One*. 2018;13(3):e0194816. doi:10.1371/journal.pone.0194816
- 278 6. Jacka B, Applegate T, Krajden M, et al. Phylogenetic clustering of hepatitis C virus  
279 among people who inject drugs in Vancouver, Canada. *Hepatology*. Nov 2014;60(5):1571-  
280 1580. doi:10.1002/hep.27310
- 281 7. Tamura K, Nei M, Kumar S. Prospects for inferring very large phylogenies by using  
282 the neighbor-joining method. *Proc Natl Acad Sci U S A*. Jul 27 2004;101(30):11030-5.  
283 doi:10.1073/pnas.0404206101
- 284 8. Rothenberg RB, Long DM, Sterk CE, et al. The Atlanta Urban Networks Study: a  
285 blueprint for endemic transmission. *AIDS*. Sep 29 2000;14(14):2191-200.  
286 doi:10.1097/00002030-200009290-00016
- 287 9. Rhodes T. Risk environments and drug harms: a social science for harm reduction  
288 approach. *Int J Drug Policy*. May 2009;20(3):193-201. doi:10.1016/j.drugpo.2008.10.003
- 289 10. van Beek I, Dwyer R, Dore GJ, Luo K, Kaldor JM. Infection with HIV and hepatitis C  
290 virus among injecting drug users in a prevention setting: retrospective cohort study. *BMJ*.  
291 Aug 15 1998;317(7156):433-7. doi:10.1136/bmj.317.7156.433
- 292 11. Hagan H, McGough JP, Thiede H, Weiss NS, Hopkins S, Alexander ER. Syringe  
293 exchange and risk of infection with hepatitis B and C viruses. *Am J Epidemiol*. Feb 01  
294 1999;149(3):203-13. doi:10.1093/oxfordjournals.aje.a009792
- 295 12. Strasfeld L, Lo Y, Netski D, Thomas DL, Klein RS. The association of hepatitis C  
296 prevalence, activity, and genotype with HIV infection in a cohort of New York City drug  
297 users. *J Acquir Immune Defic Syndr*. Jul 01 2003;33(3):356-64. doi:10.1097/00126334-  
298 200307010-00010
- 299 13. Falade-Nwulia O, Hackman J, Mehta SH, et al. Factors associated with phylogenetic  
300 clustering of hepatitis C among people who inject drugs in Baltimore. *BMC Infect Dis*. Nov 10  
301 2020;20(1):815. doi:10.1186/s12879-020-05546-x
- 302 14. Ramachandran S, Thai H, Forbi JC, et al. A large HCV transmission network enabled  
303 a fast-growing HIV outbreak in rural Indiana, 2015. *EBioMedicine*. Nov 2018;37:374-381.  
304 doi:10.1016/j.ebiom.2018.10.007
- 305 15. Sacks-Davis R, Daraganova G, Aitken C, et al. Hepatitis C virus phylogenetic  
306 clustering is associated with the social-injecting network in a cohort of people who inject  
307 drugs. *PLoS One*. 2012;7(10):e47335. doi:10.1371/journal.pone.0047335
- 308 16. Kaplan EH. Needles that kill: modeling human immunodeficiency virus transmission  
309 via shared drug injection equipment in shooting galleries. *Rev Infect Dis*. 1989 Mar-Apr  
310 1989;11(2):289-98. doi:10.1093/clinids/11.2.289
- 311 17. Tumpney M, John B, Panneer N, et al. Human Immunodeficiency Virus (HIV)  
312 Outbreak Investigation Among Persons Who Inject Drugs in Massachusetts Enhanced by  
313 HIV Sequence Data. *J Infect Dis*. 09 02 2020;222(Suppl 5):S259-S267.  
314 doi:10.1093/infdis/jiaa053

