The More You Look, the More You Find: Effects of Hepatitis C Virus Testing Interval on Reinfection Incidence and Clearance and Implications for Future Vaccine Study Design

Peter Vickerman,1,2 Jason Grebely,3 Gregory J. Dore,3,6 Rachel Sacks-Davis,4,5,6 Kimberly Page,7 David L. Thomas,8 William O. Osburn,8 Andrea L. Cox,8 Campbell K. Aitken,4,5,6 Matthew Hickman,2 and Margaret Hellard4,5,6 on behalf of the InC3 Collaborative Group

1Department of Global Health and Development, London School of Hygiene and Tropical Medicine, 2School of Social and Community Medicine, University of Bristol, United Kingdom; 3The Kirby Institute for Infection and Immunity in Society, University of New South Wales, Sydney, 4Burnet Institute, Melbourne, 5Department of Epidemiology and Preventative Medicine, Monash University, and 6Centre for Research Excellence in Injecting Drug Use, Melbourne, Australia; 7Department of Epidemiology and Biostatistics, University of California, San Francisco; and 8Department of Medicine, Johns Hopkins Medical Institution, Baltimore, Maryland

Introduction. Studies have explored whether spontaneous clearance of hepatitis C virus (HCV) infection decreases the likelihood of reinfection or increases the probability of clearance. This analysis investigates whether the conflicting findings from these studies could be due to differences in frequency of HCV RNA testing.

Methods. A model simulated the dynamics of HCV reinfection and clearance among a cohort of injection drug users. For different reinfection incidence and clearance rates, the model evaluated the accuracy of epidemiological studies that used different HCV testing frequencies.

Results. Experimental estimates for the reinfection incidence and clearance probability will be accurate (<20% error) if the testing interval is less than the reinfection clearance duration. Otherwise, experimental estimates can greatly underestimate the real values (≤66% error if reinfection duration is 1 month and the testing interval is 3 months). Uncertainty in experimental estimates also increases at lower reinfection incidences, whereas for lower clearance probabilities the uncertainty in the estimated clearance probability increases but estimated reinfection incidence decreases.

Discussion. Differences in HCV testing interval could account for most between-study variability in the estimated probability of clearing reinfections and is likely to have biased reinfection incidence estimates. Our findings suggest that a high reinfection clearance probability (>75%) is consistent with data.

Over the past 2 decades, our understanding of hepatitis C virus (HCV) has substantially improved, and pharmacologic advances have led to a rapid expansion of new and emerging therapies [1]. However, there is still no vaccine to prevent HCV infection. The development of an effective HCV vaccine has been complicated by the existence of multiple HCV genotypes, the limited availability of animal models, and complex immunological responses to HCV infection [2]. Nevertheless, given that spontaneous clearance of HCV infection occurs in about 25% of individuals [3], there is optimism that an HCV vaccine is possible.

Individuals with primary HCV infection and subsequent clearance have high-magnitude, broadly specific, and sustained cellular immune responses, while those developing persistent infection often have weaker cellular immune responses that do not persist [4, 5]. In addition, both the innate [6] and humoral immune responses (in particular, virus-specific neutralizing antibodies) [7] have been shown to be important in determining the outcome of HCV infection. Reinfec
following spontaneous HCV clearance has been observed in chimpanzees [8–13], injection drug users (IDUs) [14–22], and men who have sex with men [23, 24], demonstrating that spontaneous clearance of primary HCV infection does not provide sterilizing immunity (ie, absolute protection against a subsequent infection). However, some of these same studies among chimpanzees [8–13] and humans [17–19] show that, following HCV reinfection, there is improved control of viral replication, an attenuated course of infection, and an increased likelihood of viral clearance, compared with primary infection, suggesting that prior clearance of HCV infection may result in partial protective immunity against persistent reinfection. In chimpanzees, rapid virological control upon reinfection is associated with HCV-specific T-cell responses [10, 12, 13]. Given these data, it may be that a therapeutic vaccine that enhances spontaneous clearance of HCV infection and thereby induces partial protective immunity is more feasible than a vaccine that provides sterilizing immunity. This is consistent with other viral pathogens (eg, hepatitis B virus, human papillomavirus, influenza virus, and varicella zoster virus), in which vaccination may not provide sterilizing immunity against infection but protects against persistent infection and leads to an attenuated course of infection [25, 26].

Studies of HCV reinfection provide insight into factors important for protection against persistent HCV infection. The chimpanzee animal model has provided most data on HCV reinfection because experiments can be carefully designed to study HCV re-exposure and reinfection. However, HCV is a uniquely human disease, and recent studies of HCV reinfection in humans have provided insight into our understanding of protective HCV immunity. To explore the possible existence and nature of acquired immunity for HCV, 7 epidemiological studies over the last 10 years have explored whether prior HCV infection reduces the chance of subsequent reinfection and/or increases the likelihood of clearing any subsequent infection [14–18, 20, 22]. These studies had similar experimental designs but produced differing results (Table 1): some suggested the incidence of reinfection is less than the incidence of primary infection, others suggested it is similar, and some suggested it is greater than the incidence of primary infection. In addition, the reinfection clearance probability varies widely between the different studies, from a low of 29% [20], similar to the estimated probability of spontaneous clearance for primary infections [16], to highs of 83%–100% in more recent studies [17, 18].

These widely different findings have several plausible explanations, including such factors as small sample sizes (each study had <22 reinfections); follow-up biases, such as misclassification [27]; and differences in host and virus genotype distributions. However, one major determining factor could be differences in the average HCV RNA testing interval between studies, which varies from 1 to 16 months [28]. If some reinfections are of short duration, on the order of 1–4 months as observed in the study by Osburn et al [17], then studies employing HCV RNA testing intervals of ≥3 months are likely to miss some reinfections that clear spontaneously; this will lead to the study underestimating the incidence of reinfection and the subsequent probability of spontaneous clearance [29]. The degree to which this potentially biases the experimental results from different studies will depend on the RNA testing interval and how it compares with the likely distribution for the duration of reinfections that spontaneously clear. Indeed, if it is a substantial effect, then one would expect the probability of spontaneous clearance following reinfection and, possibly, the ratio of the incidence of HCV reinfection to the incidence of HCV primary infection to be larger in studies with shorter RNA testing intervals.

This study uses modeling to explore the degree to which between-study variations in RNA testing interval (from 1 to 12 months) could explain the variability in the experimental estimates for the probability of spontaneous clearance.
following HCV reinfection and the ratio between the incidence of primary infection and reinfection. The modeling analysis considers the implications for studies undertaken in settings with different HCV reinfection incidence levels and for a range of possible reinfection clearance probabilities.

METHODS

Data Analysis
Linear regression was used to explore the degree to which variability in the average RNA testing interval in different studies (independent variable in both models) explains the large observed variability in the different study estimates for the reinfection clearance probability (dependent variable in first model) and crude incidence rate ratio between the incidence of reinfection and primary infection (dependent variable in second model). The reinfection clearance probability for each study was weighted by the number of reinfections that occurred in that study, whereas the crude incidence rate ratio was weighted by the overall person years of risk for each study.

Model Description
A probabilistic individual-based model was developed to simulate HCV reinfection and clearance of a fixed number of IDUs over time. At first (t = 0), all IDUs in the model are assumed to have cleared their primary HCV infection, and are susceptible (state S) to reinfection. Over each discrete time step (Δt), all susceptible IDUs have a constant probability (λΔt) of becoming infected dependent on the predefined incidence of reinfection (λ), and if they become infected they are considered HCV RNA positive at the end of that time step. Once infected, an IDU either enters the infected state that leads to spontaneous clearance (state Ic) with probability (δ) or otherwise enters the infected state that leads to chronic infection (state I), which they remain in until the end of the model simulation. For those who spontaneously clear infection (Ic), this occurs after a duration τ that is either constant or sampled from a probability distribution depending on the model analysis being undertaken. Once an individual has spontaneously cleared their infection, they are assumed to become fully susceptible (state S) to reinfection again. To simulate the epidemiological studies, the model population is assumed to be closed, with no IDUs entering or leaving the cohort, and for simplicity all IDUs are assumed to have the same length of follow-up. The transitions that each IDU can undertake, with their associated probabilities or the duration in each state, are shown in Figure 1.

To simulate the results of observational studies that use different RNA testing intervals (T), the infection status of each IDU is determined at time 0, T, 2T, and so on. Changes in the infection status at consecutive time points are used to determine whether a new reinfection has occurred or a reinfection has been cleared. A new reinfection is recorded if any individual changes from susceptible to infected over 2 consecutive testing time points, whereas a reinfection clearance is recorded if any individual changes from infected to susceptible over 2 consecutive testing time points. If an individual becomes infected and always tests RNA positive at subsequent testing time points, the model assumes they did not clear their infection during the follow-up period of the experiment. Lastly, the model does not consider different HCV genotypes or include the transmission of multiple strains of HCV, which are added complexities that epidemiological studies must consider but should not affect the general insights that can be obtained from this model.

Model Analyses
For a range of “real” HCV reinfection incidence rates (λ) and reinfection clearance probabilities (δ) used as model inputs, the model was used to project what experimental estimates could be obtained from simulated epidemiological studies with differing RNA testing intervals. The model considered how these estimates would differ from the “real” values for reinfections that clear having constant but different duration or a range of durations similar to the durations observed in the study by Osburn et al [17]. The range of values considered for the “real” yearly HCV reinfection incidence, reinfection clearance probability, reinfection clearance duration, and the RNA testing interval are included in Table 2. To correspond with the small mean number of individuals followed for reinfections in each study (n = 55), all model simulations assumed a cohort of 50 individuals that are initially susceptible to reinfection and are followed for 48 months. To account for the stochastic nature of the model and because parameters were sampled across different values, 10 000 simulations were undertaken for each analysis, with each simulation using parameter values randomly sampled from the range of values allowed for each parameter shown in Table 2. These simulations were used...
to determine the median and 5th, 25th, 75th, and 95th percentiles for the experimental estimates of the HCV reinfection incidence and HCV reinfection clearance probability for each combination of the “real” HCV reinfection incidence, duration, clearance probability, and RNA testing interval.

The analysis initially focuses on the model projections for a reinfection incidence of 32 infections per 100 person-years and a 75\% clearance probability to coincide with the high reinfection incidence and clearance probability observed in the study with the shortest RNA testing interval [17]. However, following this, the model analysis considers lower reinfection incidence levels and clearance probabilities (Table 2) to explore how these parameters affect our model insights.

### RESULTS

#### Data Trends

For studies with larger average HCV RNA testing intervals, there is weak evidence for a decrease in both the estimated reinfection clearance probability (\(\beta = -.03; R^2 = 0.42; P = .12\)) and the crude incidence rate ratio (\(\beta = -.11; R^2 = 0.44; P = .10\)) between the incidence of reinfection and primary infection (Figure 2). In both regression models, the variability in the RNA testing interval between different studies accounts for about 40\% of the variability in the dependent variables.

#### Model Projections

For a real reinfection incidence of 32 infections per 100 person-years and a clearance probability of 75\%, Figure 3 shows how the experimentally estimated reinfection incidence and clearance probability varies for different durations of reinfections that clear (0.5, 1, 2, and 4 months or uniformly variable between 0.5–4 months) and for testing intervals of 1, 3, 6 or 12 months. The figure shows that the median estimated reinfection incidence and clearance probability are both likely to be very close to their real estimates if the testing interval is the same as or less than the duration of reinfections that clear (defined as “reinfection clearance duration”). However, if the testing interval is greater than the reinfection clearance duration, then the experimental estimates, especially the reinfection incidence, are likely to be greatly underestimated. For example, if the testing interval is 3 months and the reinfection clearance duration is 2 months, then the median estimated reinfection incidence will be 23 infections per 100 person-years (5th to 95th percentile range, 18–27 infections per 100 person-years), which is 28\% lower than the real reinfection incidence of 32 infections per 100 person-years, and the reinfection clearance probability will be 63\% (55%-69\%), which is 16\% lower than the real reinfection clearance probability of 75\%. For the same testing interval (3 months), similar differences are seen if the reinfection clearance duration is variable.

### Table 2. Parameter Values Used in the Model Analyses

<table>
<thead>
<tr>
<th>Model Parameter</th>
<th>Symbol</th>
<th>Values Used or Range</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Real” HCV reinfection incidence, infections per 100 person-years</td>
<td>(\lambda)</td>
<td>4, 8, 16, 32</td>
<td>Similar to range found in studies in Table 1</td>
</tr>
<tr>
<td>“Real” HCV reinfection clearance probability, %</td>
<td>(\delta)</td>
<td>25, 50, 75</td>
<td>Similar to range found in studies in Table 1</td>
</tr>
<tr>
<td>“Real” HCV reinfection duration for reinfections that clear, mo</td>
<td>(\tau)</td>
<td>0.5, 1, 2, 4; 0.5–4</td>
<td>Osburn et al’s [17] range of 1–4 mo was extended to include possibility of shorter reinfections</td>
</tr>
<tr>
<td>HCV RNA testing interval, mo</td>
<td>(T)</td>
<td>1, 3, 6, 12</td>
<td>Similar to range found in studies in Table 1</td>
</tr>
</tbody>
</table>

![Figure 2](image-url). The observed relationship between the average hepatitis C virus RNA testing interval for different studies and the estimated reinfection clearance probability (A) or crude incidence rate ratio (between the incidence of reinfection and primary infection; B) for each study. Panel A also involves comparison of the observed relationship with the model projections that assume a “real” reinfection incidence of 32 infections per 100 person-years and that 75\% of reinfections are cleared after 0.5–4 months. Error bars are 95\% confidence intervals for the data estimates and model projections. No error bars are shown in panel B because some studies did not report duration of follow-up.

HCV Testing Interval and Reinfection and Clearance • JID 2012:205 (1 May) • 1345
between 0.5 and 4 months, whereas these differences become larger if the duration of reinfection is ≤1 month (instead of 2 months) or become negligible if reinfection durations are ≥3 months. Alternatively, if the testing interval is ≥6 months and the reinfection clearance duration is ≤2 months (or variable between 0.5 and 4 months), then the estimated reinfection incidence and clearance probability could easily be less than half the real values but are likely to be very similar to the real values if the testing interval is 1 month. The only exception to this is if the reinfection clearance duration is <1 month.

Interestingly, the model projections in Figure 3 for variable reinfection clearance durations coincides well with the observed effect of different HCV RNA testing intervals on the probability of clearing a reinfection (Figure 2). This suggests that variability in the RNA testing interval could be the main driver behind variability in the reinfection clearance probabilities from different studies. It also suggests that a high reinfection clearance probability of about 75% could be consistent with the results obtained in all studies.

Figures 2 and 3 show results of analyses that considered the implications of different testing intervals for a scenario in which the reinfection incidence was 32 infections per 100 person-years and the clearance probability was 75%, both similar to the observed values from the study by Osburn et al [17]. Figure 4 shows that for both the estimated reinfection incidence and clearance probability, on average, similar relative errors are likely to occur in studies with different "real" reinfection incidence rates, but the variability in the error is much greater at lower "real" incidence levels. Conversely, for lower "real" reinfection clearance rates, the error in the estimated reinfection incidence diminishes, whereas the error in the estimated reinfection clearance probability increases.

DISCUSSION

In this study, we demonstrated quantitatively, using mathematical modeling, that studies using long HCV RNA testing intervals underestimate the incidence of HCV reinfection and the probability of spontaneous HCV clearance following reinfection. These findings provide empirical support for the hypothesis [28, 29] that long HCV RNA testing intervals may account for lower rates of reinfection among IDUs with spontaneous clearance when compared with rates of primary infection [14–18, 20, 22]. The mechanism behind this finding is simply that broad HCV testing intervals miss HCV reinfection events with spontaneous clearance, particularly if the duration of reinfection is less than the HCV RNA testing interval. The results of this study have important implications for HCV vaccine design because they suggest that, although absolute protection against primary reinfection (sterilizing immunity) is probably overestimated, the rate of spontaneous clearance of reinfection (partial protective immunity against persistent HCV reinfection) is also underestimated.

To explore the possible existence and nature of acquired immunity for HCV, 7 epidemiological studies over the last 10 years have explored whether prior HCV infection reduces the chance of subsequent reinfection and/or increases the likelihood of clearing any subsequent infection [14–18, 20, 22]. These human studies all have limitations that may explain differences in their outcomes. For example, differences in
subject characteristics, including age, ethnicity, risk behaviors, and HIV status, could have affected observed rates of primary and reinfection clearance. Crucially, the actual number of HCV exposures is unknown, and proxies such as drug use practices are incomplete at best. While chimpanzee studies count exposures, human studies only count incidents of viremia. An implication of this limitation is that long HCV RNA testing intervals underestimate exposure because reinfection and clearance events are overlooked, while viral persistence (by definition) is not (Figure 5). This can lead (Figure 5A) to an apparent predominance of viral persistence (Figure 5C) and the interpretation that protective immunity does not exist. Furthermore, the outcome of reinfection (clearance vs persistence) is often not adequately characterized because the person was not studied after reinfection was first noted, again creating bias due to overestimation of viral persistence (and lack of protective immunity). The data generated in this mathematical modeling study provide quantitative evidence to support this model of HCV infection.

Studies in chimpanzees [8–13] and humans [17, 19] have demonstrated that, following HCV reinfection, there is improved control of viral replication, an attenuated course of infection, and an increased likelihood of viral clearance, compared with primary infection, suggesting that prior clearance of an HCV infection may result in partial protective immunity against persistent reinfection. This is supported by the findings of this study. Given these data, it may be that a therapeutic vaccine that enhances spontaneous clearance of HCV infection, thereby inducing partial protective immunity, is more feasible than a vaccine that provides sterilizing immunity. However, further research needs to confirm this finding, to understand what factors confer this protection and to determine whether the clearance rate for primary infection may have been similarly underestimated among IDUs in epidemiological studies.

As with all modeling analyses, the findings of this analysis are limited because they are not based on empirical evidence but on a model’s representation of reality. The model includes several simplifying assumptions that could have affected its projections. First, the model simulates the transmission of a single HCV strain and does not consider multiple or super infections. Further, it is not known whether spontaneous clearance of HCV reinfection differs by heterologous exposure to different HCV genotypes, so the issue of cross-genotype protection is not addressed in this model. Second, we assumed a constant probability of infection among IDUs over time, with the model ignoring any heterogeneity in injection risk. Although these simplifications may affect whether the model can replicate the detailed results from an individual cohort, especially in modeling the difference between reinfection and primary incidence, they should not affect the general insights of the model on the effect of testing interval on experimental accuracy. Third, the model only considers reinfections with duration between 0.5 and 4 months, whereas it is possible that some reinfections may be shorter than half a month, although data are lacking. This means that the model can only evaluate whether differences in study results could be due to different testing intervals but cannot estimate the likely error in the reinfection clearance probability or incidence estimate for the study with the shortest testing interval of 1 month [17].

Figure 4. The relative error in the experimental estimates of the hepatitis C virus reinfection incidence and clearance probability for different “real” reinfection incidences (A) or clearance probabilities (B). All projections assume a duration of reinfection of 0.5–4 months and a testing interval of 3 months, panel A assumes a real reinfection clearance probability of 75%, and panel B assumes the real reinfection incidence of 32 infections per 100 person-years. Median is middle line, 25th/75th percentiles are limits of boxes and 5th/95th percentiles are whiskers.
Although this study suggests that differences in HCV RNA testing interval could explain much of the apparent differences between the findings of existing studies, it is likely that other factors may have played a role, such as differences in HCV genotype distribution and host susceptibility, sex, HIV coinfection, or confounding and follow-up biases. First, if primary infections are more likely among people with greater injection risk behavior, then in the absence of any immune protection, one would expect the reinfection incidence to be elevated compared with the incidence of primary infection. In contrast, if higher risk IDUs are more likely to be lost to follow-up, then this could have the opposite effect of resulting in a lower reinfection incidence than the incidence of primary infection. It was impossible to test these hypotheses because the studies did not have comparable measures of injection risk. Second, if the probability of spontaneous clearance varies by

![Figure 5. Impact of sampling frequency on perceptions of hepatitis C virus (HCV) reinfection outcome. A, Hypothetical virological profile for a participant with initial spontaneous HCV clearance, several HCV reinfection events with reduced duration, and peak HCV viremia and subsequent spontaneous clearance of reinfection in each instance. B, Hypothetical virological profile for a participant with initial spontaneous HCV clearance, several HCV reinfection events with similar duration, and peak HCV viremia and subsequent spontaneous clearance of reinfection in each instance. C, Hypothetical virological profile for a participant with spontaneous HCV clearance and HCV reinfection with persistent viremia. Solid lines indicate HCV RNA levels (light grey shading indicates persistent viremia). HCV RNA assessment time points are indicated by vertical dotted lines. Legends indicate the hypothetical frequency of HCV RNA testing (in weeks), the number of viremic events that would be identified with such a hypothetical HCV RNA testing frequency, and the apparent outcome of infection based on the last available time point for that HCV RNA testing frequency (reproduced with permission from Grebely et al [29]).]
host and/or viral genotype [30] or sex [3], then the probability of spontaneous clearance could vary by setting and be elevated among those individuals that clear their primary infection. In addition, HIV coinfection could also reduce the probability of spontaneous clearance, as suggested by 0 of 11 HIV-infected individuals clearing their HCV reinfection in the studies included in Table 1, and so settings in which the HIV prevalence is higher could have a lower probability of spontaneous clearance. Unfortunately, the small number of reinfection cases (90) described in the studies included here (and elsewhere) prevented any reliable assessment of the importance of these factors. However, although it is possible that they may have affected the findings of the studies to some extent, they are unlikely to have diminished the considerable effect that variations in testing interval may have had on the results.

The findings in this study illustrate the substantial error that could exist in experimental estimates of reinfection incidence and clearance for epidemiological studies that use testing intervals of >3 months. A major implication of this work is that future studies evaluating clearance of HCV reinfection in humans as a measure of protective immunity should have short HCV RNA testing intervals (eg, 1 or 2 months) to accurately capture reinfection events. Ideally, these studies would include detailed virological and immunological components. Unfortunately, more frequent follow-up is expensive and also raises logistical challenges in the recruitment and follow-up of individuals at a high risk of infection (eg, IDUs). However, there are cohorts that have been successful in this regard [17]. This work also has implications for testing other infectious diseases, such as chlamydial infection and gonorrhea, in which higher rates of infection and reinfection have been observed when more frequent testing among high-risk groups has been implemented [31].

In the near future, it is unlikely that an HCV vaccine will be developed that confers sterilizing immunity against HCV, defined as absolute protection against reinfection. However, the primary goal of vaccination would be to prevent chronic hepatitis C, which causes cirrhosis and hepatocellular cancer. As such, a vaccine protecting against the development of HCV persistence may be more feasible. Our results provide further encouragement to the quest for an effective HCV vaccine.

Notes

Financial support. P. V. is funded by a UK Medical Research Council New Investigators Award (G0701627). G. D. is supported by a National Health and Medical Research Council (NHMRC) Practitioner Fellowship. J. G. is funded by a NHMRC Career Development Fellowship. R. S. D. is funded by a NHMRC postgraduate scholarship and a Centre for Research Excellence in Injecting Drug Use top-up scholarship, and M. Hellard is supported by a NHMRC Senior Research Fellowship. C. A. is funded by the Burnet Institute, W. O. and A. C. are funded by the US National Institutes of Health (NIH; grant U19AI088791), D. T. is funded by DAR01016078 and K. P. is supported by the NIH (grants 5R01DA016017 and 1R01DA031056-01A1) and the University of California, San Francisco, Liver Center (P30 DK026743).

Potential conflicts of interest. All authors: No reported conflicts. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References