Hepatitis C virus infection spontaneous clearance: Has it been underestimated?

Houssein H. Ayoub[a,b,e,**], Hiam Chemaitelly[a], Ryosuke Omori[c,d], Laith J. Abu-Raddad[a,b,*]

[a] Infectious Disease Epidemiology Group, Weill Cornell Medicine – Qatar, Cornell University, Qatar Foundation – Education City, Doha, Qatar
[b] Department of Healthcare Policy and Research, Weill Cornell Medicine, Cornell University, New York, USA
[c] Division of Bioinformatics, Research Center for Zoonosis Control, Hokkaido University, Sapporo, Hokkaido, Japan
[d] JST, PRESTO, 4-1-8 Honcho, Kawaguchi, Saitama, 332-0012, Japan
[e] Department of Mathematics, Statistics, and Physics, Qatar University, Doha, Qatar

ARTICLE INFO

ABSTRACT

Objectives: Hepatitis C virus (HCV) clearance rate ($f_{clearance}$) is defined as the proportion of infected persons who will spontaneously clear their infection after acute infection. We aimed to estimate $f_{clearance}$ using a novel approach that avoids limitations in existing estimates, and to clarify the link between $f_{clearance}$ and HCV viremic rate—the latter being the proportion of RNA positivity among those antibody positive.
Methods: A mathematical model was developed to describe HCV transmission. $f_{clearance}$ was estimated by fitting the model to probability-based and nationally representative population-based data for Egypt (Egypt 2008 and Egypt 2015) and USA (NHANES A and NHANES B). Uncertainty and sensitivity analyses were conducted.
Results: $f_{clearance}$ was estimated at 39.9% (95% uncertainty interval (UI): 34.3%–46.4%) and 33.5% (95% UI: 29.2%–38.3%) for Egypt 2008 and Egypt 2015 data, respectively; and at 29.6% (23.0%–37.1%) and 39.9% (31.2%–51.0%) for NHANES A and NHANES B data, respectively. $f_{clearance}$ was found related to HCV viremic rate through (approximately) the formula $f_{clearance} = 1.16\,(1 - \text{HCV viremic rate})$. HCV viremic rate was higher with higher risk of HCV exposure. Robustness of results was demonstrated in uncertainty and sensitivity analyses.
Conclusion: One-third of HCV-infected persons clear their infection spontaneously, higher than earlier estimates—the immune-system capacity to clear HCV infection may have been underestimated.

© 2018 The Author(s). Published by Elsevier Ltd on behalf of International Society for Infectious Diseases. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Hepatitis C virus (HCV), first identified in 1989 (Choo et al., 1989; Kuo et al., 1989), is a major cause of liver fibrosis, cirrhosis, and cancer (Hajarizadeh et al., 2013). It is estimated that 1–3% of the population in most countries are infected with HCV (Mohd Hanafiah et al., 2013; Lavanchy, 2011). A key highlight of the

natural history of HCV infection is that a proportion of infected persons spontaneously clear the infection after enduring acute infection (Hajarizadeh et al., 2013). Earlier studies, following the discovery of the virus, suggested that only 10–15% of infected persons clear their infection (Di Bisceglie, 2000). Having an accurate and precise estimate of HCV clearance rate is important for HCV response, and will inform HCV treatment in early infection, planning of health services provision, and estimation and projections of HCV chronic infections and disease burden (Micallef et al., 2006; Grebely et al., 2014).

With the interest in clarifying HCV natural history, a number of longitudinal cohort studies were conducted and reported their findings in recent years (Micallef et al., 2006; Grebely et al., 2014; Amin et al., 2007; Seeff, 2002; Soriano et al., 2008). Although different studies have had different lengths of follow-up, making comparability less certain, they estimated a broad range for HCV clearance of 0%–57% within 6–48 months after infection (Micallef et al., 2006; Grebely et al., 2014; Amin et al., 2007; Seeff, 2002;

Soriano et al., 2008)—the best estimate was about 25% after an acute-infection duration of 16.5 weeks (Micallef et al., 2006; Grebely et al., 2014). These studies have also indicated that infection clearance is predicated on a complex interplay of host and virus factors (Hajarizadeh et al., 2013) such as female sex (Micallef et al., 2006; Grebely et al., 2014; Page et al., 2009), *IL28B* CC genotype (Grebely et al., 2014), and HCV genotype 1 (Grebely et al., 2014), among others (Ray et al., 1999; Li and Lemon, 2013; Lemon, 2010; Takaki et al., 2000).

While cohort studies provide a direct approach to estimating HCV clearance rate, this approach suffers from methodological limitations and potential biases (Micallef et al., 2006; Amin et al., 2007). The heterogeneity across studies in design, study population, sample size, inclusion criteria, length of follow-up duration, and analysis method contributed to a broad range of estimates and added complexity to the estimates' interpretation and comparability (Amin et al., 2007; Micallef et al., 2007). Reported clearance rates were often based on small samples of less than 50 participants and follow-up durations of less than 12 months, though evidence suggests that undetectable RNA levels could be reached over longer time intervals of up to 48 months (Amin et al., 2007; Micallef et al., 2007). Several studies opted for retrospective testing of stored HCV RNA samples for the assessment of clearance, but there are known limitations for samples stored under suboptimal conditions (Amin et al., 2007; Micallef et al., 2007).

The recruitment of participants may have been a key limitation in these studies and poses a question about their representativeness and generalizability. Studies tended to recruit individuals who were symptomatic and viremic at baseline (Micallef et al., 2006; Amin et al., 2007), but HCV infection is mostly asymptomatic (Hajarizadeh et al., 2013)—only few studies assessed clearance rate among sero-incident cases (Micallef et al., 2006; Page et al., 2009). Most studies recruited opportunistically from populations such as people who inject drugs (PWID), or post-transfusion HCV patients. The risk of reinfection in these populations is often high, and this could have influenced the identification of a unique acute infection (Micallef et al., 2006). PWID, a marginalized population by definition, may not have been also representative of the socio-demographic characteristics, prevailing HCV genotypes, or mode of acquisition and associated inoculum effect, of incident infections in the population at large (Hajarizadeh et al., 2013; Grebely et al., 2014; Pawlotsky et al., 1995; Messina et al., 2015). For example, females have a higher clearance rate than males (Grebely et al., 2014), but are often under-represented in the predominantly male PWID populations (Mumtaz et al., 2014). The difficulty in detecting asymptomatic acute infections, and the lack of a reliable diagnostic test to differentiate between acute and chronic infection stages may have as well affected the validity of clearance rate estimates, and their representativeness of clearance in the population at large (Amin et al., 2007).

Another critical limitation in existing studies relates to the heterogeneity in clearance definition in terms of using a single HCV Ribonucleic Acid (RNA) negative test, as opposed to two negative tests within a period of one to six months (Micallef et al., 2006; Grebely et al., 2014; Amin et al., 2007). It has been further shown that even slight variations in case definition and analysis method can underestimate clearance rate (Micallef et al., 2006), and generate estimates ranging from 14% to 68%, on the same primary data (Amin et al., 2007).

In light of these largely unavoidable limitations in cohort studies, there is a need to provide an estimate of clearance rate using a methodology that is independent of these limitations. We present here a novel approach to estimate HCV *clearance rate* ($f_{clearance}$)—defined as the proportion of HCV infected persons who spontaneously clear their infection after the acute stage of infection. This approach also clarifies the subtle link between

$f_{clearance}$ and HCV *viremic rate*, with the latter being defined as the proportion of individuals who are HCV antibody (Ab) positive and HCV RNA positive, out of all who are HCV Ab positive regardless of RNA status, as measured in a given cross-sectional survey. Of note is that $f_{clearance}$ and HCV viremic rate are technically not rates, but are strictly *proportions* with no time unit, and therefore should not in principle be labelled as rates. However, to avoid confusion with prevailing convention, we labeled them accordingly for consistency with existing literature, where these two measures have been labelled as "rates".

The fundamental concept of the present approach is that HCV natural history effects at the individual level manifest themselves at the population level. Starting from probability-based and nationally representative population-level data, we used mathematical modeling and simulations of the infection process to estimate the average HCV clearance rate in the population.

Though this approach relies on an indirect method, the strength of this analysis lies in its independence from most of the biases and limitations affecting current empirical measures. The approach also capitalizes on the availability of quality population-based data, and provides an estimate that is representative of the socio-demographic and mode of acquisition diversity in the population at large. In a sense, this approach is effectively equivalent to conducting a cohort study for the entire population of a country, but *in silico*.

Materials and methods

Conceptual framework and key definitions

We used a conceptual framework for HCV natural history based on current knowledge of HCV biology (Figure 1) (Hajarizadeh et al., 2013). Infected persons are assumed to develop primary acute infection. A fraction ($f_{clearance}$) of these individuals clear their infection spontaneously after the acute stage of infection, while the rest ($1 - f_{clearance}$) become chronically infected and positive for both HCV Ab and HCV RNA. For those who clear their infection, they can be re-infected, and thus go through a secondary acute infection stage. A proportion of the latter individuals will clear the infection while the rest become chronically infected.

To avoid confusion of subtle and closely related measures, Figure 1 includes also the exact definitions used in the present study for HCV clearance rate and HCV viremic rate.

Key definitions

➤ **HCV clearance rate ($f_{clearance}$):** proportion of HCV infected persons who spontaneously clear their infection

➤ **HCV viremic rate:** proportion of individuals who are HCV antibody (Ab) positive and HCV RNA positive, out of all who are HCV Ab positive, as measured in a given cross-sectional survey.
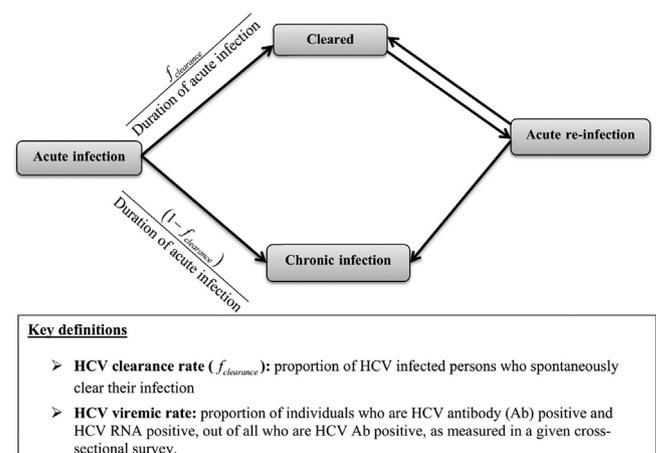
Figure 1. Conceptual framework and key definitions for hepatitis C virus (HCV) natural history.

Model description

A deterministic compartmental model was developed to describe HCV transmission in the population at large (Figure S1 in Supplementary material (SM)). The model is an extension of previous models (Vickerman et al., 2007; Vickerman et al., 2012; Deuffic-Burban and Yazdanpanah, 2013; Ayoub and Abu-Raddad, 2017; Ayoub et al., 2018), and consists of a system of coupled nonlinear differential equations that stratify the population according to HCV status and stage of infection, and level of risk of exposure (SM). To account for heterogeneity in risk of exposure in the population, the model incorporates 15 risk groups covering a spectrum of HCV exposure risk, from low (say blood donors or pregnant women) to high risk, with the three highest risk groups in the model representing PWID.

Informed by earlier work (Ayoub and Abu-Raddad, 2017; Ayoub et al., 2018; Handcock and Jones, 2004; Hamilton et al., 2008; Cuadros et al., 2011; Omori et al., 2015), we assumed that the distribution of the population across risk groups follows a gamma distribution (SM). Individuals who leave their risk group are distributed proportionally across all risk groups. The mixing between individuals in the different risk groups is determined by a mixing matrix (SM) that allows a range of mixing behaviors varying from fully assortative (mixing only with individuals in the same risk group) to fully proportionate (mixing with individuals with no preferential bias for any specific risk group) (Garnett and Anderson, 1993; Awad and Abu-Raddad, 2014).

Risk of infection was captured by the force of infection that factors in the effective number of contacts conducive for HCV transmission, HCV transmission probability, and mixing among risk groups (SM). We assumed that the distribution of risk of exposure across the different risk groups follows a power law function (SM), as suggested by earlier modeling analyses (Ayoub and Abu-Raddad, 2017; Ayoub et al., 2018; Awad and Abu-Raddad, 2014; Barendregt et al., 2003; Barrat et al., 2004). Further details on the model structure can be found in SM.

Model parameterization

We parameterized our model using current HCV natural history and transmission data (Table S1 in SM). The model was applied to two countries—Egypt and the United States of America (USA)—where HCV biomarker data for the population at large has been collected through rigorous probability-based and nationally representative surveys. The Egypt data were derived from the 2008 and 2015 Egypt Demographic and Health Surveys (EDHS) (El-Zanaty and Way, 2009; Ministry of Health and Population [Egypt] et al., 2015). HCV Ab prevalence was estimated at 14.7% (95% confidence interval (CI): 14.0–15.4%) in 2008 and at 10.0% (95% CI: 9.3–10.7%) in 2015. HCV viremic rate was assessed at 67.1% (95% CI: 64.5–69.6%) in 2008 and at 70.2% (95% CI: 67.2–73.1%) in 2015. For ease of reference, these two data sources have been labeled thereafter as Egypt 2008 and Egypt 2015, respectively.

The USA data were derived from the continuous series of the National Health and Nutrition Examination Surveys (NHANES) (NHANES, 1999–2012NHANES, 1999NHANES, 1999–2012). HCV Ab prevalence and HCV viremic rate were estimated for each NHANES round between 1999 and 2012 (before the recent scale up in treatment), and then pooled using DerSimonian-Laird random-effects meta-analyses (DerSimonian and Laird, 1986). Pooled HCV Ab prevalence was estimated at 1.4% (95% CI: 1.3–1.6%), and pooled HCV viremic rate was estimated at 74.0% (95% CI: 57.9–79.6%). This data source has been labeled thereafter as NHANES A.

Since NHANES laboratory procedures entail the additional testing of individuals with undetermined HCV Ab results for HCV RNA positivity, we also derived a second pooled estimate for the NHANES viremic rate including as denominator both individuals confirmed as HCV Ab positive and those with an undetermined HCV Ab status. This yielded a viremic rate of 64.6% (95% CI: 56.2–72.6%). This data source has been labeled thereafter as NHANES B.

The model was additionally applied to six countries where HCV biomarker data were available through population-based surveys, but without sufficient description of the methodology to assess whether the data where collected using strictly probability-based and nationally representative sampling of standardized and comparable methodology to that of the EDHS and NHANES surveys. These included surveys from Brazil (Pereira et al., 2013), India (Lee et al., 2014), Ireland (Thornton et al., 2012), Latvia (Tolmane et al., 2011), Netherlands (Vriend et al., 2012), and Vietnam (Do et al., 2014).

Data on the population proportion of PWID, a key determinant of HCV epidemiology across countries, were retrieved from global and regional reviews (Mumtaz et al., 2014; Aceijas et al., 2004; Lansky et al., 2014).

Model fitting

We estimated the population $f_{clearance}$ by fitting model output to HCV Ab prevalence, HCV viremic rate, and the population proportion of PWID for each country, assuming endemic equilibrium. Model fits were implemented in MATLAB® (MATLAB®, 2013) using a nonlinear least-square fitting method based on the Nelder-Mead simplex algorithm, described in Lagarias et al. (Lagarias et al., 1998).

Uncertainty analysis

We conducted multivariable uncertainty analyses to determine the range of uncertainty around our estimates for $f_{clearance}$ by varying the model parameters. We specifically implemented 500 runs of the model applying at each run Monte Carlo sampling from uniform probability distributions that were generated from either the CI, or assuming (if uncertainty is not captured by CI) ±25% uncertainty around the point estimates of these parameters. In each run, the parameters' values were randomly selected from their specified ranges, and the model was refitted to data.

The parameters included in the uncertainty analyses are HCV transmission probability for the different stages of infection, duration of each infection stage, proportion of individuals in secondary acute infection who clear their infection spontaneously, HCV viremic rate in each setting, duration that an individual spends in a specific risk group, the degree of assortativeness in the mixing, the scale parameter in the gamma distribution of the population across risk groups, and the exponent parameter in the power law function of the distribution of risk of exposure. The mean of the resulting distribution for $f_{clearance}$ and its associated 95% uncertainty interval (UI) were derived.

Sensitivity analyses

Several sensitivity analyses were conducted using the Egypt 2008 model as an illustrative example. These analyses assessed the robustness of our predictions for $f_{clearance}$ to variations in key measures or parameters that may not be known with sufficient precision, or suspected to potentially influence our predictions. These measures or parameters included HCV Ab prevalence in the population, statistical distribution of the duration of primary acute infection, statistical distribution of the population across the 15 risk groups, risk of exposure variation among the 15 risk groups in the model, and fraction of individuals in secondary acute infection who clear their infection spontaneously.
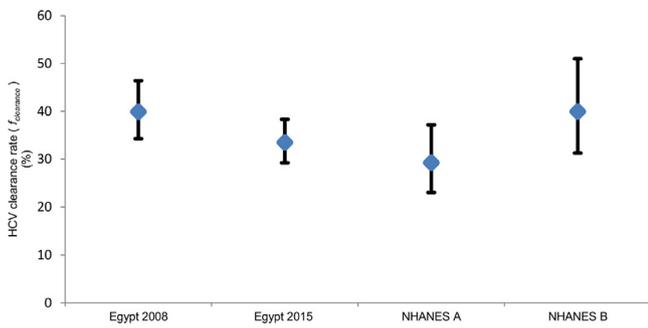
Figure 2. Estimated mean and 95% uncertainty interval (UI) for hepatitis C virus (HCV) clearance rate ($f_{clearance}$). The estimates were conducted based on data from the 2008 and 2015 Egypt Demographic and Health Surveys (EDHS) and the continuous series of the National Health and Nutrition Examination Surveys (NHANES).

We further conducted two additional sensitivity analyses. In the first analysis, we estimated $f_{clearance}$ in Egypt by fitting the model output to the actual temporal variation in HCV Ab prevalence, that is by relaxing the assumption of endemic equilibrium. In the second sensitivity analysis, using the Egypt 2008 model as an illustrative example, we assessed the robustness of our estimate for $f_{clearance}$ to large variations in the uncertainty of the input parameters by allowing 50% uncertainty around the point estimates, instead of only 25%.

Results

Robust model fits were obtained for the fitted epidemiological measures in all country-specific estimations (Table S2 in SM). In Egypt, $f_{clearance}$ was estimated at 39.9% (95% UI: 34.3%–46.4%) for Egypt 2008, and at 33.5% (95% UI: 29.2%–38.3%) for Egypt 2015 data (Figure 2). In the USA, $f_{clearance}$ was estimated at 29.6% (23.0%–37.1%) for NHANES A, and at 39.9% (31.2%–51.0%) for NHANES B data (Figure 2).

Figure 3 shows the estimated $f_{clearance}$ versus HCV viremic rate in the population using (in addition) the survey data from Brazil (Pereira et al., 2013), India (Lee et al., 2014), Ireland (Thornton et al., 2012), Latvia (Tolmane et al., 2011), Netherlands (Vriend et al., 2012), and Vietnam (Do et al., 2014). Figure 3 also shows the functional relationship between the estimated $f_{clearance}$ and HCV

viremic rate, as generated using the Egypt 2008 model as an illustrative example. $f_{clearance}$ was found to linearly depend on HCV viremic rate. The best fit line to the model predictions yielded the following linear relationship:

$$f_{clearance} = 1.16(1 - \text{HCV viremic rate from population}$$
$$-\text{based survey}) \qquad (1)$$

All country-specific estimates for $f_{clearance}$, from each individual country-specific model, were found to lie very close to the curve generated using the Egypt 2008 model (by changing the background population antibody prevalence).

The above result establishes the functional relationship between $f_{clearance}$ and HCV viremic rate in the whole population, but HCV viremic rate is most often measured for specific risk populations, such as PWID, rather than for the whole population using probability-based and nationally representative surveys. Figure 4 shows the estimated HCV viremic rate for each risk group of the modeled 15 risk groups in the population, using the Egypt 2008 and NAHNES A models as illustrative examples. The figure further displays HCV viremic rate for the whole population and indicates $f_{clearance}$ as estimated above for each of Egypt 2008 and NHANES A.

HCV viremic rate was found to increase with the level of risk of exposure to HCV infection, and was highest among groups at highest risk, particularly PWID (Figure 4). In the higher risk groups, HCV viremic rate can be substantially higher than the HCV viremic rate in the whole population—it could not be used to estimate $f_{clearance}$ using Eq. (1). Meanwhile, HCV viremic rate in the lower risk populations (say blood donors or pregnant women) was found to be closer to HCV viremic rate in the whole population (with slight underestimation)—it could be used to estimate approximately $f_{clearance}$ using Eq. (1).

The five conducted sensitivity analyses assessing the robustness of our predictions to variations in the key measures or parameters affirmed the robustness of the estimates for $f_{clearance}$ (Figure S2). They also demonstrated the impact of variations in the different input parameters on the estimated $f_{clearance}$. Only small variations were observed in the estimated $f_{clearance}$ despite extreme variations in the measures and parameters that may not be known with sufficient precision, or suspected to potentially influence our predictions.

In the additional sensitivity analysis assessing the impact of relaxing the assumption of endemic equilibrium, the estimated
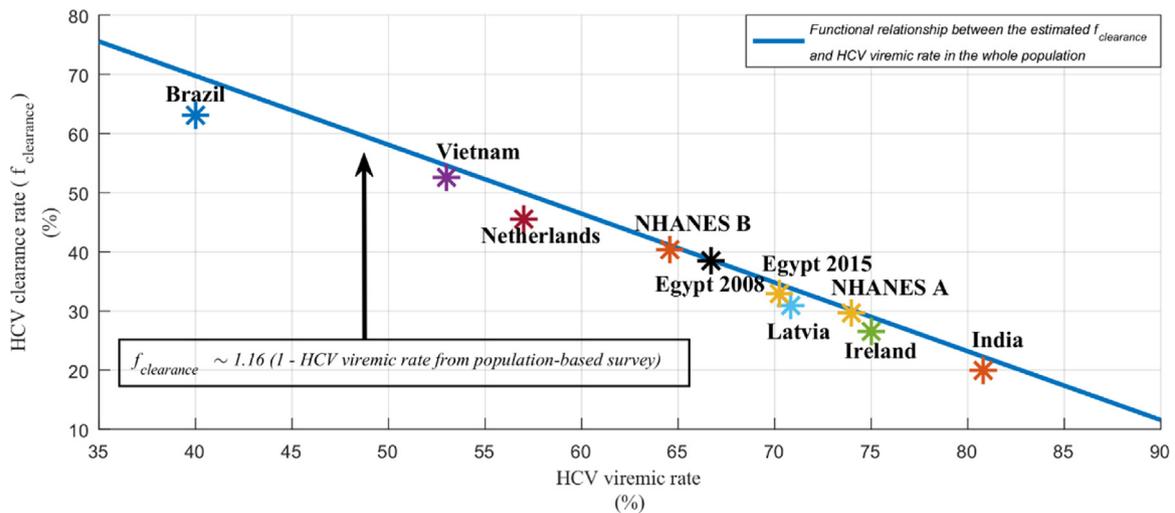

Figure 3. Relationship between hepatitis C virus (HCV) clearance rate ($f_{clearance}$) and HCV viremic rate in the whole population. The figure shows (solid line) the functional relationship between the estimated $f_{clearance}$ and HCV viremic rate in the whole population, as generated using the Egypt 2008 model (by varying background antibody prevalence in the population), as an illustrative example. The figure shows also $f_{clearance}$ versus HCV viremic rate in the whole population as estimated using 10 different population-based survey data.
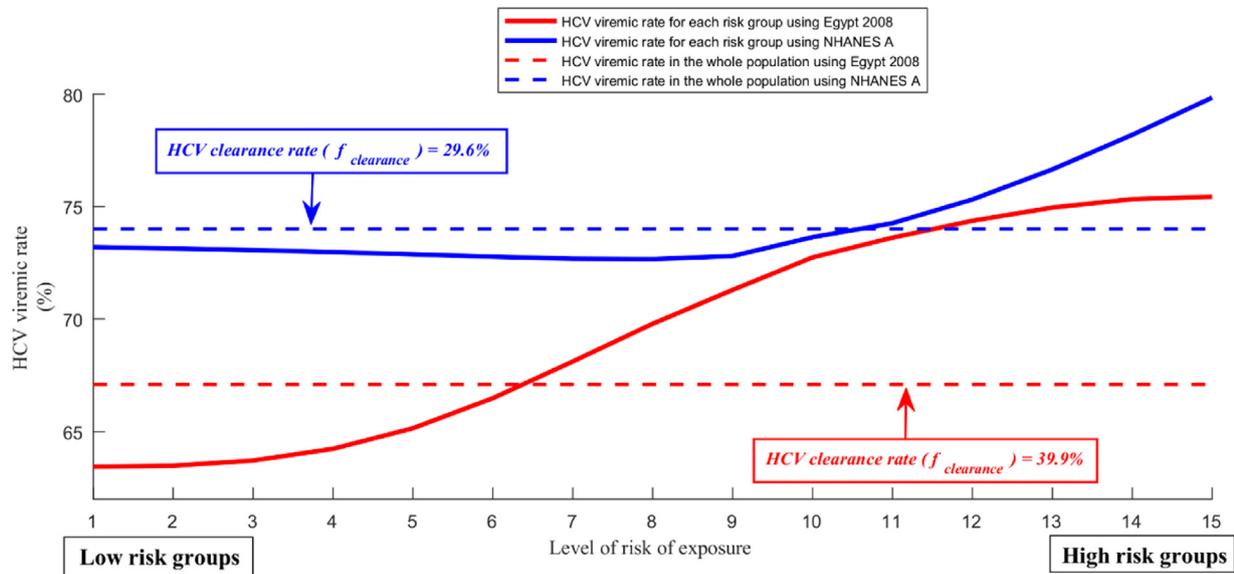
Figure 4. Hepatitis C virus (HCV) viremic rate for each HCV risk group of the modeled 15 risk groups in the population, using the Egypt 2008 and NAHNES A models as illustrative examples. The figure also shows HCV viremic rate for the whole population.

$f_{clearance}$ for Egypt was 37.5%, within the range of the estimates assuming endemic equilibrium, thereby affirming our conclusions. In the sensitivity analysis assessing the impact of large uncertainty in the input parameters, $f_{clearance}$ for Egypt 2008 was estimated at 39.3% (95% UI: 33.4%–46.1%)—affirming also our conclusions despite the larger uncertainty interval.

Discussion

We estimated HCV clearance rate ($f_{clearance}$) using a novel analytical approach that avoids most limitations of the estimates of longitudinal cohort studies. We found that $f_{clearance}$ is in the range of 30–40%—substantially higher than the best estimates of cohort studies of about 25% (Micallef et al., 2006; Grebely et al., 2014). We also derived a simple relationship that can predict $f_{clearance}$ in the whole population from a measure of HCV viremic rate in this population (Eq. (1)). While quality probability-based and nationally representative measures of HCV viremic rate may not be available for most countries, we found that HCV viremic rate in low risk populations, such as blood donors or pregnant women, can approximately represent HCV viremic rate in the whole population, and could be used to estimate $f_{clearance}$ for this population. Meanwhile, we found that HCV viremic rate in high risk populations, such as PWID, may not be representative of that in the whole population and should not be used to infer $f_{clearance}$ in the population.

Our finding that existing empirical measures from longitudinal cohort studies are likely to have underestimated HCV spontaneous clearance should not be surprising. Cohort studies reported inconsistent estimates that ranged between 0%–57% (Micallef et al., 2006; Grebely et al., 2014; Amin et al., 2007; Seeff, 2002). A broad range of limitations have also been identified in these studies, as reviewed in the introduction section above, and these limitations could have affected these studies, as well as their generalizability and representativeness.

While we introduced a method for estimating $f_{clearance}$ from HCV viremic rate, existing literature suggests large variations in measured HCV viremic rate (Harfouche et al., 2017). Figure 3 demonstrates this variability for the 10 population-based measures of HCV viremic rate from different countries. While it is conceivable that there could be true variations in HCV viremic rate

reflecting underlying heterogeneity in the factors that affect spontaneous clearance such as the population proportion of PWID, type of population affected by HCV infection, gender, circulating genotype, and level of HIV coinfection, we are more inclined to believe that this variability arises mainly from differences in the complex laboratory methods used to assess HCV viremic rate.

Assessment of HCV viremic rate requires a two-test algorithm, for HCV Ab and HCV RNA, and the diagnostic assays and protocols can vary from one study to another. Even small variation in analysis methods, such as definition of the denominator in HCV viremic rate, can lead to large variations in the estimated rate. An example to this end can be seen in the difference between the two estimates of NHANES A and NHANES B, where the difference relates solely to how the denominator was defined. Another example is the variation in the viremic rate in NHANES in recent rounds versus earlier rounds, which in part may be due to a minor change in the testing protocol starting from 2013 (Centers for Disease Control and Prevention, 2016, 2018a, 2018b).

A recent comprehensive systematic meta-analysis of HCV viremic rate measures in diverse populations in the Middle East and North Africa supports this conjecture (Harfouche et al., 2017). Though HCV viremic rate was found to vary extensively across studies, the pooled means were similar irrespective of country or subregion, population HCV Ab prevalence, or study sampling method, among other factors. The overall pooled mean of all 178 measures was estimated at 67.6% (95% CI: 64.9–70.3%), closely similar to the probability-based and nationally representative estimates of the EDHS and NHANES (El-Zanaty and Way, 2009; Ministry of Health and Population [Egypt] et al., 2015; NHANES, 1999–2012NHANES, 1999NHANES, 1999–2012). These considerations advocate for use of one standardized methodology for the definition and measurement of HCV viremic rate in the literature. This will not only facilitate estimation of $f_{clearance}$ in a given population, but (importantly) it will facilitate the use of the viremic rate for assessing and monitoring the scale up and coverage of HCV treatment programs, as we progress towards the target of HCV elimination by 2030 (World Health Organization, 2016).

Our study has limitations. The introduced approach for estimating $f_{clearance}$ is best implemented on probability-based and nationally representative data, but such data are not yet

available in most countries. In absence of a standardized methodology for estimating HCV viremic rate, measures will tend to vary widely curtailing the utility of the introduced method to estimate $f_{clearance}$. A pooled measure (over many studies) for the viremic rate, instead of a study-specific measure, may alleviate this challenge as has been shown recently in the systematic meta-analysis of HCV viremic rate data (Harfouche et al., 2017). We used an elaborate mathematical model to capture the complexity of HCV dynamics, but some of our model assumptions may not hold and this could have affected our predictions. However, our model yielded robust fits to the epidemiological measures (Table S2 in SM), and we performed multiple sensitivity analyses on the model output to assess the robustness of our estimates—these analyses affirmed our findings (Figure S2).

We did not incorporate treatment coverage for HCV chronically infected individuals, but treatment coverage has been very low up to the recent scale up of direct-acting antivirals (DAAs) programs (Ayoub and Abu-Raddad, 2017; World Health Organization, 2016; Polaris Observatory, 2018; Egypt Ministry of Health and Population, 2014), and therefore could not have affected the model input data, such as those of EDHS and NHANES, or the model output. We assumed that those exposed to HCV infection will be HCV Ab positive for life, which may have underestimated the clearance rate if considerable fraction of individuals lose their antibodies over time. We assessed the clearance rate in Egypt and USA where the dominant genotypes are 4 (Mahmud et al., 2018) and 1 (Klevens et al., 2012), respectively, but these results may not be generalizable to other settings where other genotypes, such as genotype 3 (Messina et al., 2015), are dominant. For simplicity, we did not explicitly incorporate gender or full dynamics details of injecting drug use in the model.

Our study has key strengths. We used a novel approach to estimate $f_{clearance}$ that builds on the success of related approaches applied for HIV infection (Bellan et al., 2015; Chemaitelly et al., 2014) among other infections (Omori et al., 2018). Our approach provided an independent estimate for $f_{clearance}$ that avoids most limitations of empirical measures. $f_{clearance}$ estimation was based on state of the art population-based input data generated through rigorous probability-based and nationally representative sampling. The $f_{clearance}$ estimates were also generated as averages for the whole population, and therefore are representative of the diversity of factors that exists in the whole population (and may affect clearance) such as socio-demography, mode of acquisition and associated inoculum effect, heterogeneity in risk of exposure, interplay of host and virus factors, infection symptoms, and risk of reinfection. This is to be contrasted with cohort studies that recruit from specific settings and populations, such as PWID, who may not be representative of the wider population exposed to HCV infection.

Moreover, existing literature (Micallef et al., 2006; Grebely et al., 2014; Amin et al., 2007; Harfouche et al., 2017) supports our modeling results that the viremic rate is higher among PWID than among the wider population. For example, the viremic rate in the NHANES A analysis was 74.9% (67.3–81.9%) among PWID versus 69.9% (59.2–79.7%) in the wider population—while the difference is not statistically significant, it is suggestive of higher viremic rate among PWID.

Importantly, we introduced a simple equation for deriving $f_{clearance}$ from HCV viremic rate, and this equation clarifies the subtle link between $f_{clearance}$ and HCV viremic rate, two distinct measures that seem often to be confused as one measure. The $f_{clearance}$ estimate was found higher than estimated using the simplistic formula $f_{clearance} = (1 - \text{HCV viremic rate})$. The difference is due to reinfection, as people who clear their infection can be re-exposed leading to chronic infection despite earlier clearance, as well as to the finite and not too small months-long duration of primary acute infection.

Lastly, we conducted an extensive uncertainty analysis around the estimated clearance rates. This analysis captured more sources of uncertainty than possible through conventional maximum likelihood parameter-estimation methods. For example, the estimated clearance rates for Egypt 2008 and NHANES A using a maximum likelihood approach are 39.9% (95% CI: 39.8–40.3%) and 29.6% (95% CI: 29.3–29.8%), respectively—the narrow confidence intervals being a consequence of the very large sample sizes of these national surveys.

In conclusion, about one-third of HCV infected persons clear their infection spontaneously, indicating that empirical measures from longitudinal cohort studies may have underestimated the capacity of the host immune system to clear HCV infection. This finding has implications for the ongoing efforts to estimate HCV infection burden, and to plan for health services provision. This finding also may have implications for our understanding of the biological determinants of HCV spontaneous clearance. It may hint that a strategy for HCV vaccine development could be a vaccine that does not necessarily prevent infection, but modulates immune response towards conditions that increase the capacity of the host immune system to clear HCV infection spontaneously.

Conflict of interest

The authors have no conflicts of interest to disclose.

Authors' contributions

HHA conducted the mathematical modeling analyses and co-wrote the first draft of the paper. HC supported the model parameterization through statistical analyses and co-wrote the first draft of the paper. RO contributed to the modeling analyses. LJA conceived and led the design of the study, analyses, and drafting of the article. All authors have read and approved the final manuscript.

Disclose funding received for this work

Others.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ijid.2018.07.013.

References

Aceijas C, Stimson GV, Hickman M, Rhodes T. Global overview of injecting drug use and HIV infection among injecting drug users. AIDS 2004;18:2295–303.

Amin J, Law M, Micallef J, Jauncey M, Van Beek I, Kaldor J, et al. Potential biases in estimates of hepatitis C RNA clearance in newly acquired hepatitis C infection among a cohort of injecting drug users. Epidemiol Infect 2007;135:144–50.

Awad SF, Abu-Raddad LJ. Could there have been substantial declines in sexual risk behavior across sub-Saharan Africa in the mid-1990s?. Epidemics 2014;8:9–17.

Ayoub HH, Abu-Raddad LJ. Impact of treatment on hepatitis C virus transmission and incidence in Egypt: a case for treatment as prevention. J Viral Hepat 2017;24:486–95.

Ayoub HH, Al-Kanaani Z, Abu-Raddad LJ. Characterizing the temporal evolution of the hepatitis C virus epidemic in Pakistan. J Viral Hepat 2018;25:670–9.

Barendregt JJ, Van Oortmarssen GJ, Vos T, Murray CJ. A generic model for the assessment of disease epidemiology: the computational basis of DisMod II. Popul Health Metr 2003;1:4.

Barrat A, Barthelemy M, Pastor-Satorras R, Vespignani A. The architecture of complex weighted networks. Proc Natl Acad Sci U S A 2004;101:3747–52.

Bellan SE, Dushoff J, Galvani AP, Meyers LA. Reassessment of HIV-1 acute phase infectivity: accounting for heterogeneity and study design with simulated cohorts. PLoS Med 2015;12:e1001801.

Centers for Disease Control and Prevention. National Health and Nutrition Examination Survey: 2015–2016 Data Documentation, Codebook, and Frequencies. https://wwwn.cdc.gov/Nchs/Nhanes/2015-2016/HEPC_I.htm. Published September 2017. [Accessed 10 July 2018].

Centers for Disease Control and Prevention. National Health and Nutrition Examination Survey: 2007–2012 Data Documentation, Codebook, and Frequencies. https://wwwn.cdc.gov/Nchs/Nhanes/2007-2008/SSHCV_E.htm. Published February 2015. [Accessed 10 July 2018].

Centers for Disease Control and Prevention. National Health and Nutrition Examination Survey: 2013–2014 Data Documentation, Codebook, and Frequencies. https://www.cdc.gov/Nchs/Nhanes/2013-2014/SSHEPC_H.htm. Published January 2018. [Accessed 10 July 2018].

Chemaitelly H, Awad SF, Abu-Raddad LJ. The risk of HIV transmission within HIV-1 sero-discordant couples appears to vary across sub-Saharan Africa. Epidemics 2014;6:1–9.

Choo Q-L, Kuo G, Weiner AJ, Overby LR, Bradley DW, Houghton M. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. Science 1989;244:359–62.

Cuadros DF, Crowley PH, Augustine B, Stewart SL, García-Ramos G. Effect of variable transmission rate on the dynamics of HIV in sub-Saharan Africa. BMC Infect Dis 2011;11:1.

DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177–88.

Deuffic-Burban S, Yazdanpanah Y. Mathematical modeling: a tool for selecting agents with complementary modes of action. J Hepatol 2013;59:1346–8.

Di Bisceglie AM. Natural history of hepatitis C: its impact on clinical management. Hepatology 2000;31:1014–8.

Do SH, Yamada H, Fujimoto M, Ohisa M, Matsuo J, Akita T, et al. High prevalences of hepatitis B and C virus infections among adults living in Binh Thuan province, Vietnam. Hepatol Res 2014;45:259–68.

Egypt Ministry of Health and Population. Plan of Action for the Prevention, Care & Treatment of Viral Hepatitis, Egypt 2014–2018. http://www.emro.who.int/images/stories/egypt/VH_Plan_of_Action_FINAL_PRINT1.pdf. 2014.

El-Zanaty F, Way A. Egypt demographic and health survey 2008. Cairo, Egypt: Ministry of Health, El-Zanaty and Associates, and Macro International; 2009.

Garnett GP, Anderson RM. Factors controlling the spread of HIV in heterosexual communities in developing countries: patterns of mixing between different age and sexual activity classes. Philos Trans R Soc B Biol Sci 1993;342:137–59.

Grebely J, Page K, Sacks-Davis R, Loeff MS, Rice TM, Bruneau J, et al. The effects of female sex, viral genotype, and IL28B genotype on spontaneous clearance of acute hepatitis C virus infection. Hepatology 2014;59:109–20.

Hajarizadeh B, Grebely J, Dore GJ. Epidemiology and natural history of HCV infection. Nat Rev Gastroenterol Hepatol 2013;10:553–62.

Hamilton DT, Handcock MS, Morris M. Degree distributions in sexual networks: a framework for evaluating evidence. Sex Transm Dis 2008;35:30.

Handcock MS, Jones JH. Likelihood-based inference for stochastic models of sexual network formation. Theor Popul Biol 2004;65:413–22.

Harfouche M, Chemaitelly H, Kouyoumjian SP, Mahmud S, Chaabna K, Al-Kanaani Z, et al. Hepatitis C virus viremic rate in the Middle East and North Africa: systematic synthesis, meta-analyses, and meta-regressions. PLoS One 2017;12: e0187177.

Klevens RM, Hu DJ, Jiles R, Holmberg SD. Evolving epidemiology of hepatitis C virus in the United States. Clin Infect Dis 2012;55:S3–9.

Kuo G, Choo Q, Alter H, Gitnick G, Redeker A, Purcell R, et al. An assay for circulating antibodies to a major etiologic virus of human non-A, non-B hepatitis. Science 1989;244:362–4.

Lagarias JC, Reeds JA, Wright MH, Wright PE. Convergence properties of the Nelder–Mead simplex method in low dimensions. SIAM J Optim 1998;9:112–47.

Lansky A, Finlayson T, Johnson C, Holtzman D, Wejnert C, Mitsch A, et al. Estimating the number of persons who inject drugs in the United States by meta-analysis to calculate national rates of HIV and hepatitis C virus infections. PLoS One 2014;9: e97596.

Lavanchy D. Evolving epidemiology of hepatitis C virus. Clin Microbiol Infect 2011;17:107–15.

Lee M-H, Yang H-I, Yuan Y, L'Italien G, Chen C-J. Epidemiology and natural history of hepatitis C virus infection. World J Gastroenterol 2014;20:9270.

Lemon SM. Induction and evasion of innate antiviral responses by hepatitis C virus. J Biol Chem 2010;285:22741–7.

Li K, Lemon SM. Innate immune responses in hepatitis C virus infection. Semin Immunopathol 2013;35:53–72.

MATLAB®. The language of technical computing. The MathWorks, Inc.; 2013.

Mahmud S, Al-Kanaani Z, Chemaitelly H, Chaabna K, Kouyoumjian SP, Abu-Raddad LJ. Hepatitis C virus genotypes in the Middle East and North Africa: distribution, diversity, and patterns. J Med Virol 2018;90:131–41.

Messina JP, Humphreys I, Flaxman A, Brown A, Cooke GS, Pybus OG, et al. Global distribution and prevalence of hepatitis C virus genotypes. Hepatology 2015;61:77–87.

Micallef J, Kaldor J, Dore G. Spontaneous viral clearance following acute hepatitis C infection: a systematic review of longitudinal studies. J Viral Hepat 2006;13:34–41.

Micallef JM, Macdonald V, Jauncey M, Amin J, Rawlinson W, van Beek I, et al. High incidence of hepatitis C virus reinfection within a cohort of injecting drug users. J Viral Hepat 2007;14:413–8.

Ministry of Health and Population [Egypt], El-Zanaty and Associates [Egypt], ICF International. Egypt health issues survey 2015. 2015 Cairo, Egypt and Rockville, Maryland, USA, p. 234.

Mohd Hanafiah K, Groeger J, Flaxman AD, Wiersma ST. Global epidemiology of hepatitis C virus infection: new estimates of age-specific antibody to HCV seroprevalence. Hepatology 2013;57:1333–42.

Mumtaz GR, Weiss HA, Thomas SL, Riome S, Setayesh H, Riedner G, et al. HIV among people who inject drugs in the Middle East and North Africa: systematic review and data synthesis. PLoS Med 2014;11.

NHANES. National Health and Nutrition Examination Survey. http://www.cdc.gov/nchs/nhanes/nhanes_questionnaires.htm; 1999–2012.

Omori R, Chemaitelly H, Abu-Raddad LJ. Dynamics of non-cohabiting sex partnering in sub-Saharan Africa: a modelling study with implications for HIV transmission. Sex Transm Infect 2015;91:451–7.

Omori R, Chemaitelly H, Althaus CL, Abu-Raddad LJ. Does infection with *Chlamydia trachomatis* induce long-lasting partial immunity? Insights from mathematical modelling. Sex Transm Infect 2018; in press.

Page K, Hahn JA, Evans J, Shiboski S, Lum P, Delwart E, et al. Acute hepatitis C virus infection in young adult injection drug users: a prospective study of incident infection, resolution, and reinfection. J Infect Dis 2009;200:1216–26.

Pawlotsky J-M, Tsakiris L, Roudot-Thoraval F, Pellet C, Stuyver L, Duval J, et al. Relationship between hepatitis C virus genotypes and sources of infection in patients with chronic hepatitis C. J Infect Dis 1995;171:1607–10.

Pereira LM, Martelli CM, Moreira RC, Merchan-Hamman E, Stein AT, Cardoso RMA, et al. Prevalence and risk factors of hepatitis C virus infection in Brazil, 2005 through 2009: a cross-sectional study. BMC Infect Dis 2013;13:60.

Polaris Observatory. http://polarisobservatory.org/polaris/hepC.htm. 2018.

Ray SC, Wang YM, Laeyendecker O, Ticehurst JR, Villano SA, Thomas DL. Acute hepatitis C virus structural gene sequences as predictors of persistent viremia: hypervariable region 1 as a decoy. J Virol 1999;73:2938–46.

Seeff LB. Natural history of chronic hepatitis C. Hepatology 2002;36:S35–46.

Soriano V, Mocroft A, Rockstroh J, Ledergerber B, Knysz B, Chaplinskas S, et al. Spontaneous viral clearance, viral load, and genotype distribution of hepatitis C virus (HCV) in HIV-infected patients with anti-HCV antibodies in Europe. J Infect Dis 2008;198:1337–44.

Takaki A, Wiese M, Maertens G, Depla E, Seifert U, Liebetrau A, et al. Cellular immune responses persist and humoral responses decrease two decades after recovery from a single-source outbreak of hepatitis C. Nat Med 2000;6:578–82.

Thornton L, Murphy N, Jones L, Connell J, Dooley S, Gavin S, et al. Determination of the burden of hepatitis C virus infection in Ireland. Epidemiol Infect 2012;140:1461–8.

Tolmane I, Rozentale B, Keiss J, Arsa F, Brigis G, Zvaigzne A. The prevalence of viral hepatitis C in Latvia: a population-based study. Medicina 2011;47:532–5.

Vickerman P, Hickman M, Judd A. Modelling the impact on hepatitis C transmission of reducing syringe sharing: London case study. Int J Epidemiol 2007;36:396–405.

Vickerman P, Martin NK, Hickman M. Understanding the trends in HIV and hepatitis C prevalence amongst injecting drug users in different settings—implications for intervention impact. Drug Alcohol Depend 2012;123:122–31.

Vriend HJ, de Coul ELO, Van De Laar TJ, Urbanus AT, Van Der Klis FR, Boot HJ. Hepatitis C virus seroprevalence in The Netherlands. Eur J Public Health 2012;22:819–21.

World Health Organization. Combating hepatitis B and C to reach elimination by 2030. Geneva, Switzerland: World Health Organization; 2016.