Published in final edited form as:

Contemp Clin Trials. 2008 September ; 29(5): 671–678. doi:10.1016/j.cct.2008.03.006.

Evaluation of operational chronic infection endpoints for HCV vaccine trials

Minhee Kang, Ph.D.¹ and Uwe Nicolay, Ph.D.²

¹Harvard School of Public Health, Boston, MA, USA

²Novartis Vaccines, Marburg, Germany

Abstract

Hepatitis C virus (HCV) is a leading cause of chronic liver disease. The natural history of HCV infection is heterogeneous, and a person infected with HCV can clear the virus or progress to a chronic infection. The chronic infection can remain asymptomatic for decades before the development of liver cirrhosis and/or carcinoma. Currently, there are no assays that can differentiate a transient infection (an acute infection that would clear) from a chronic infection, and serial HCV RNA testing is used to operationally define chronic hepatitis C (e.g. detectable HCV over 6 months). Therefore, HCV vaccine trial planning can benefit from the assessment of the endpoint candidates that are aimed at the chronic infection. Operationally defined endpoints based on the virological tests at study visits have been previously studied in the context of human papillomavirus (HPV) vaccine trials. However, HCV natural history is different from HPV, requiring separate considerations. In this work, several definitions of chronic infection that are based on the periodically observed HCV RNA statuses are evaluated, using a multi-state, time-homogenous Markov model for transient and chronic infections under various infection settings. Our results show some inflation in the type I error in the log-rank test on the vaccine efficacy against chronic infections in the presence of vaccine efficacy related to transient infections. A type I error up to almost four times the planned rate of 5% is observed in one setting. Overall, simple operational endpoints yield higher power than more complex endpoints, but the simplest endpoint is most affected by the type I error inflation and misclassification error due to the assay imperfection.

Keywords

Log-rank test; Multi-state process; HCV; Vaccine efficacy

1. Introduction

Infection with hepatitis C virus (HCV) is a worldwide problem that has affected approximately 170 million persons [1] and nearly 4 million in the United States [2]. An acute HCV infection may be cleared by the host immune system (transient infection) or may result in a chronic, persistent infection, which over years can lead to cirrhosis and hepatocellular carcinoma. An important feature of HCV natural history is the high rate of chronic infections in the magnitude

Address: Department of Biostatistics, 655 Huntington Ave, Boston, MA 02115, Phone: 617-432-2819, Fax: 617-432-3163, Email: mkang@hsph.harvard.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

of 75–85% [3]. Clinical reviews have quoted estimates of viral clearance as 25% [4,5] or lower [6]. The time to clearance has been described to range from 3 to 24 months [7]. The HCV clearance estimates have ranged widely, likely due to the characteristics of the populations studied and evaluation methods based on the available data. The definitions of chronic infection also vary among studies. Potential biases in these estimates have been discussed in [8]. Six major HCV genotypes have been identified, with the highest prevalence of about 80% for genotype 1 among the HCV infected persons in the US [9]. The annual incidence rate can be as high as 20% or higher among the intravenous drug users [10,11]. The current standard of treatment is least effective for individuals with chronic HCV infections of genotype 1, with clearance in 40–50% [12,13]. Hence, a vaccine that prevents chronic HCV infections, particularly of genotype 1, will be an important public health contribution, and there are HCV vaccines under development [14–16].

HCV is often referred to as the silent killer, because the infected person can remain asymptomatic for decades before developing cirrhosis and/or hepatocellular carcinoma [9]. With such lengthy asymptomatic period, consideration of a clinical endpoint for an HCV vaccine trial is not practical. However, a difficulty in choosing an endpoint that relies on HCV detection is that currently there is no laboratory test to determine if the HCV detection will lead to a chronic infection or a transient infection that will eventually clear. Two types of laboratory assays are available for HCV. Enzyme immunoassays are commonly used as a screening tool to detect antibodies to HCV, but they cannot distinguish between a current infection and an infection that the individual has recovered from. Polymerase chain reaction (PCR) assays detect the virus in the individual's blood to assess current infection status [13], and chronic infection is sometimes defined by the persistence of HCV in the blood for at least 6 months [17].

A successful vaccine needs to show efficacy against chronic infections, and a vaccine that mainly prevents the transient infection would be considered suboptimal. In this respect, the HCV vaccine design considerations share similar issues with those previously considered in human papillomavirus (HPV) vaccine trial designs, where the infection preceding the disease is not sufficient for progression to the disease (since it may be transient). In this paper, we consider several operational definitions of HCV infection that may be used to indicate chronic infections, in the context of a vaccine trial. These definitions are operationally defined as testing positive for HCV detection in consecutive visits, similar to the definitions considered in the HPV setting [18]. We aim to evaluate various operational definitions of HCV chronic infection as endpoint candidates for a vaccine trial by examining the type I error rate and the power of the log-rank test using each endpoint candidate. We also consider possible effects of infection status diagnostic test errors (misclassification errors) on the performance of the candidate endpoints.

2. Underlying Process

For the evaluation of the infection endpoints, we assume the 4-state, time-homogeneous Markov process $X(\bullet)$, depicted in Figure 1, for the development of transient HCV infection that clears without clinical intervention (state 2a) and chronic infection (state 2b). In the HCV application, an individual may be HCV negative ($X(t)=1$), infected with HCV that is transient ($X(t)=2a$), or infected with chronic HCV infection ($X(t)=2b$) that eventually presents clinical symptoms ($X(t)=3$). An individual may acquire and clear an infection of the transient type repeatedly, but a chronic HCV infection is irreversible on its path to development of liver disease. If the process can be observed for a long time, the symptomatic state 3, that follows state 2b, may be observed.

Let $p_{jk}(s,t)$ represent the probability that an individual in state j at time s is in state k at time t , where $j,k=1,2a,2b$ and $s \leq t$. For example, $p_{11}(s,t)$ denotes the probability that an individual who is HCV negative at time s is also HCV negative at time t . Since the process is assumed to be time-homogeneous, $p_{jk}(s,t) = p_{jk}(0,t-s)$. The intensity function for transition from state j to state k , or cause-specific hazard at time t , is denoted by λ_{ij} and defined as,

$$\lambda_{jk} = \lim_{\delta \downarrow 0} \frac{p_{jk}(t, t+\delta) - p_{jk}(t, t)}{\delta}, j \neq k.$$

Let $P(s,t)$ and Λ denote the 3×3 matrices of transition probabilities and intensities, respectively, where the j th diagonal element (λ_{jj}) is the negative of the rate of leaving state j : $\lambda_{jj} = -\sum_{k,k \neq j} \lambda_{jk}$. The relationship between the transition probabilities and the transition rates is given by (c.f. [19]):

$$P(s,t) = \exp^{(t-s)\Lambda} = \sum_{m=0}^{\infty} \frac{\Lambda^m (t-s)^m}{m!}. \quad (1)$$

The time that the process stays in a state before making a transition to a different state is exponentially distributed, with the mean sojourn time in state j given by $-1/\lambda_{jj}$.

3. Event definitions

3.1 Operational endpoints based on observations

Let $0 = v_0 < v_1 < v_2 \dots < v_M$ denote the pre-specified visit times where an individual is evaluated for the presence of HCV infection by the HCV PCR assay. Let Y_m denote the observed state at time v_m , $m=0,1,2,\dots,M$, where $Y_m=2$ if HCV is detected, and $Y_m=1$ if not. Note that in the observed test results, states 2a and 2b are not distinguishable, hence state 2 is used to represent the observable positive test, whether it be transient or chronic.

We consider four operational definitions to capture the chronic infection event, based on the observed diagnostic test results:

- a single positive test (+)
- positive tests at two consecutive visits (++)
- positive tests at three consecutive visits (+++)
- positive tests at three consecutive visits not followed by a negative test in the follow-up (++...+).

The first three endpoints are similar to the ones considered in [18] in the application to HPV. In a study with 3-month visit intervals, the endpoints +++ and ++...+ correspond to chronic infection marked by HCV persistence for at least 6 months found in the literature [17]. The last endpoint is unusual in that it requires the individual to be observed as positive until the end of the scheduled follow-up in the study, assumed to be the same for all subjects in our consideration. Due to the imperfect nature of the chronic infection definition, it is not clear which of these - or another - operational definition would be preferable in the vaccine trial setting. The purpose of this paper is to evaluate these definitions as candidate endpoints in a vaccine clinical trial that aims to prove efficacy against chronic infections. These operational definitions correspond to the following observation patterns.

$$\begin{aligned} +: & (Y_m=1, Y_{m+1}=2) \\ ++: & (Y_m=1, Y_{m+1}=2, Y_{m+2}=2) \\ +++: & (Y_m=1, Y_{m+1}=2, Y_{m+2}=2, Y_{m+3}=2) \\ ++ \dots +: & (Y_m=1, Y_{m+1}=2, Y_{m+2}=2, \dots, Y_M=2) \end{aligned}$$

The corresponding endpoint time for the first endpoint is v_{m+1} for some $m=0,1,2,\dots$. For the second and third endpoints, the event times may be determined rather arbitrarily as v_{m+2} and v_{m+3} , respectively, at the time the endpoint criteria are met, or retrospectively as v_{m+1} for both, when HCV is first detected. We take the latter approach, as is commonly done. In truth, the infection would have occurred between the visits v_m and v_{m+1} , if there are no HCV detection errors. The last endpoint is determined at the end of the individual follow-up, and the infection time is considered to be when the first positive diagnostic test is observed, at v_{m+1} .

3.2 Link to Markov process and misclassification error

The HCV PCR assay cannot distinguish a transient (state 2a) from a chronic infection (state 2b). Hence, in the case of chronic infection, state 2 would be observed at each succeeding visit, if there were no misclassification errors in HCV detection and if the study duration is too short to observe state 3. Let X_m denote the true observable state of the process at the visit time v_m ; that is $X_m=X(v_m)$, where $X_m \in \{1,2\}$. We allow X_m to be subject to misclassification errors due to the imperfect HCV detection assay, where an individual's infection status is subject to error. Let the random variable Y_m introduced in 3.1 to be the observed value at time v_m , which will equal X_m , if and only if the assay is correct in HCV detection. As an example, presence of HCV below what can be detected by the assay due to the assay lower limit of detection may contribute to the detection error. We make the following conditional independence assumption regarding the misclassification errors due to imperfect HCV tests, as assumed in [18]:

$$\Pr(Y_0, Y_1, K, Y_M \mid X_0, X_1, K, X_M) = \prod_{m=0}^M \Pr(Y_m \mid X_m),$$

where X_m is the true infection status and Y_m is the status subject to misclassification at v_m . That is, conditioned on the true values of $X(\bullet)$ at the visit times, the distribution of Y_m depends only on the value of $X(\bullet)$ at v_m . These error probabilities are given by,

$$\gamma_{lk} = \Pr(Y_m = k \mid X_m = l). \quad (2)$$

To better understand the differences among the true underlying process ($X(\bullet)$, $X_m=1,2a,2b$), the true observable process ($X(\bullet)$, $X_m=1,2$) and the observed process with misclassification ($Y(\bullet)$, $Y_m=1,2$) with an example, consider the endpoint ++ observed at v_{m+1} and v_{m+2} , i.e. $Y_m=1$, $Y_{m+1}=2$ and $Y_{m+2}=2$. A number of event histories can lead to this observation. Here are only a few possibilities, denoting the true states at the visit times v_{m+1} and v_{m+2} in boldface and the unobserved states between the visit times in regular font.

- Chronic infection between v_m and v_{m+1} , without misclassification errors: **1-2b-2b**.
- Transient infection between v_m and v_{m+1} , then clearance and recurrence between v_{m+1} and v_{m+2} that clears after v_{m+2} , without misclassification errors: **1-2a-1-2a-1**.
- Transient infection between v_m and v_{m+1} , then clearance between v_{m+1} and v_{m+2} , with misclassification at v_{m+2} : **1-2a-1**.

Some of these events may be rare. For instance, the second event is highly unlikely in the process where clearance time is long relative to the visit interval and infection rate is low.

4. Hypotheses, type I error and power

4.1 Vaccine effects and hypotheses of interest

In our evaluation of the various candidate endpoints, we consider the settings where a vaccine may affect the rate of chronic infection ($\lambda_{1,2b}$), the rate of transient infection ($\lambda_{1,2a}$) and/or the clearance rate of the transient infection ($\lambda_{2a,1}$). The vaccine efficacies in prevention of

persistent, chronic infection (VE_P), in prevention of transient infection (VE_T) and in clearing transient infection (VE_C) are defined as,

$$VE_P = 1 - \frac{\lambda_{1,2b}^v}{\lambda_{1,2b}^p}, VE_T = 1 - \frac{\lambda_{1,2a}^v}{\lambda_{1,2a}^p} \text{ and } VE_C = 1 - \frac{\lambda_{2a,1}^p}{\lambda_{2a,1}^v},$$

where λ_{jk}^v and λ_{jk}^p are the transition intensities in the vaccine and the placebo groups, respectively.

Although our interest is in testing for VE_P , the potential efficacies against acquiring and in clearing transient infections, given by VE_T and VE_C , may distort the type I error of the test on VE_P , as discussed in [18]. While the global null hypothesis,

$$H_0^*: VE_P = 0, VE_T = 0, VE_C = 0,$$

maintains the type I error given by the nominal size of the test, the composite null hypothesis of interest,

$$H_0: VE_P = 0, VE_T = \eta, VE_C = \delta,$$

that the vaccine has no effect on the transition intensity for developing a chronic infection may not. When H_0 holds but not H_0^* , the test for vaccine efficacy may have a distorted size, and this study investigates the distortions in the HCV application by considering various values for η and δ .

4.2 Type I error and power

The null hypothesis of interest H_0 is composite. When the null hypothesis is composite, the type I error can be treated as power [18]. The type I error for the composite null hypothesis of interest H_0 is the probability of obtaining a significant result when $VE_P = 0$ but at least one of δ or η is nonzero; that is, the type I error for the test on VE_P depends on the values of VE_T and VE_C . This is given by the power of the test where the null hypothesis is the global null hypothesis H_0^* and the alternative hypothesis is the composite null hypothesis H_0 . The power of the test of H_0 depends on the values of VE_T and VE_C , as well as the value of VE_P .

We consider the log-rank test for the inference on vaccine efficacy in preventing chronic infections, VE_P , a commonly used statistical test in an efficacy trial. The time until the occurrence of an event, defined as one of the four considered in Section 3, will have a distribution that depends functionally on the parameters $\lambda_{1,2a}$, $\lambda_{1,2b}$, $\lambda_{2a,1}$, γ_{lk} and the number and times of scheduled visits. We apply an approximation to the power of the log-rank test derived in [18] for a two-sided test,

$$\Phi \left(Z_{\alpha/2} + \frac{\sqrt{N} \sum_{m=1}^M |\pi_{mv} - \pi_{mp}|}{\sqrt{\sum_{m=1}^M (\pi_{mv} + \pi_{mp})}} \right),$$

where $\Phi(z)$ denotes the cumulative normal distribution, N denotes the equal sample size **in each** of the vaccine and placebo groups, α denotes the type I error rate, and π_{mv} and π_{mp} denote the probabilities of observing an event at time v_m in the vaccine and placebo groups, respectively. The expressions for the event probabilities π_{mj} , $j=v,p$, are determined by enumerating all possible paths and misclassification outcomes that correspond to the event occurring at time v_m . These event probabilities depend on $\lambda_{1,2a}$, $\lambda_{1,2b}$, $\lambda_{2a,1}$ as stated in equation (1) and on γ_{lk} , defined in equation (2).

To illustrate the calculation of event probabilities, π_{mv} and π_{mp} , consider the endpoint ++...+ at visit $M-4$ when there are $M=8$ post-vaccination visits. This is given by the observed sequence, $\{Y_0=1, Y_1=1, Y_2=1, Y_3=1, Y_4=2, Y_5=2, Y_6=2, Y_7=2, Y_8=2\}$, where Y_0 is the last vaccination visit. The probability of meeting the endpoint ++...+ at this visit (denoted here as $\pi_{4j}, j=v, p$) is given by:

$$\begin{aligned} & \Pr(Y_0=1, Y_1=1, Y_2=1, Y_3=1, Y_4=2, Y_5=2, Y_6=2, Y_7=2, Y_8=2) \\ & + \Pr(Y_0=1, Y_1=2, Y_2=1, Y_3=1, Y_4=2, Y_5=2, Y_6=2, Y_7=2, Y_8=2) \\ & + \Pr(Y_0=1, Y_1=1, Y_2=2, Y_3=1, Y_4=2, Y_5=2, Y_6=2, Y_7=2, Y_8=2) \\ & + \Pr(Y_0=1, Y_1=2, Y_2=2, Y_3=1, Y_4=2, Y_5=2, Y_6=2, Y_7=2, Y_8=2). \end{aligned}$$

Each of the above probabilities can be calculated by,

$$\begin{aligned} & \Pr(Y_0=y_0, Y_1=y_1, Y_2=y_2, Y_3=y_3, Y_4=y_4, Y_5=y_5, Y_6=y_6, Y_7=y_7, Y_8=y_8) \\ & = \sum_x \gamma_{y_0, x_0} \gamma_{y_1, x_1} \gamma_{y_2, x_2} \gamma_{y_3, x_3} \gamma_{y_4, x_4} \gamma_{y_5, x_5} \gamma_{y_6, x_6} \gamma_{y_7, x_7} \gamma_{y_8, x_8} \Pr(X=x), \end{aligned}$$

where $x=(x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8)$ is every possible sequence of unobservable states consisting of 1, 2a, 2b that may produce the given observed sequence. Then, by the Markov property,

$$\begin{aligned} \Pr(X=x) &= \Pr(X(v_8)=x_8 | X(v_7)=x_7) \Pr(X(v_7)=x_7 | X(v_6)=x_6) \Pr(X(v_6)=x_6 | X(v_5)=x_5) \\ & \wedge \Pr(X(v_2)=x_2 | X(v_1)=x_1) \Pr(X(v_1)=x_1 | X(v_0)=x_0) \Pr(X(v_0)=x_0) \\ & = p_{x_7, x_8}^j(v_7, v_8) p_{x_6, x_7}^j(v_6, v_7) p_{x_5, x_6}^j(v_5, v_6) \wedge \Pr(X(v_0)=x_0), \end{aligned}$$

where $p_{x_m, x_{m+1}}^j(v_m, v_{m+1})$, for $j=v, p$, are the transition probabilities and $\Pr(X(v_0)=x_0)$ is the initial probability. This initial probability would be 1 in studies that only include subjects who are not detected with HCV at the initial follow-up. Expressions for the probabilities of observing the event at other time points or of observing other operational definitions of the event at a particular time point can be calculated similarly.

5. Model parameters

5.1 Transition intensities, vaccine efficacies and misclassification errors

The HCV sero-conversion rate in the intravenous drug user population has been reported to vary between 10 and 20 per 100 person-years [10,20–22], although a rate as high as 37 per 100 person-years has been reported [11]. We assume that 80% of HCV infections are of genotype 1 [9]. We consider three genotype 1 infection hazard rates ($\lambda_{1,2a} + \lambda_{1,2b}$), 0.00702, 0.01084 and 0.01488 per month, and two proportions of transient infection, 0.15 and 0.35 [9] to derive 6 settings for $\lambda_{1,2a}$ and $\lambda_{1,2b}$. The clearance times varies widely in the literature, and we consider two mean clearance times for transient infections of 6 and 24 months [7,17,23], corresponding to $\lambda_{2a,1}=0.167$ and 0.0417, respectively. The vaccine efficacy rates of 0%, 25% and 65% are assumed for VE_T and VE_C , and the rate of 65% is assumed for VE_p . For the misclassification errors, we consider the HCV detection assay specificity error of 5%; that is, 5% error in incorrectly testing positive when the individual is not infected. The HCV PCR assay has high sensitivity (correctly testing positive when the individual is infected) [24,25], and perfect sensitivity is assumed for this study.

5.2 Sample size and study length

We fix the study sample size in our study to evaluate the type I error and power for the four operationally defined endpoints in various settings. The methods in Freedman [26] for a 2-sided log-rank test at a type I error rate of 5% are used to derive the sample size. The genotype 1 infection hazard rate of 0.0108 per month and the chronic infection proportion of 0.65 are used: $\lambda_{1,2a}=0.00379$ and $\lambda_{1,2b}=0.00704$. Assuming the vaccine efficacy of 65%, 34 events are

required for 80% power. The individual study duration is assumed to be 30 months, where the vaccination series ends at 6 months and the follow-up infection status observations are at months 9, 12, 15, 18, 21, 24, 27 and 30 (corresponding to $v_1, v_2, \dots, v_8$). Because the candidate endpoint events are infection statuses over time (except for +), the follow-up time duration when the event can be counted depends on the endpoint. For example, the endpoint +++ is observed over 3 visits. With our visit schedule, this can only be observed at 6 visit times: at months 9, 12, ..., 24 counting the event time when HCV is first detected. In contrast, the endpoint + can be observed at 8 visit times, at month 9 or later. Approximating the follow-up time of 18 months in the Freedman method, corresponding to the ++ endpoint, yields $N=210$ in each group.

6. Results

6.1 Type I errors in testing $VE_P > 0$ when $VE_T > 0$ and/or $VE_C > 0$

Table 1A, Table 2A and Table 3A pre N denotes the equal sample size in each of the vaccine and placebo groups sent the type I errors for the genotype 1 infection rates 0.00702, 0.0108 and 0.0149 per month, respectively. We first examine the setting of no misclassification errors. When the vaccine efficacies for all three modes are assumed to be zero, the type I error rates are maintained at 5%, as expected. However, when VE_T and/or VE_C is greater than zero, the type I errors are distorted. The type I error increases as VE_T and/or VE_C increases and as the proportion of transient infection increases from 15% to 35%. As expected, the endpoint + performs the poorest, with the type I error over 10% in the setting of 35% transient infection proportion and high VE_T of 65%, and this endpoint is most influenced by VE_T . In general, the type I error increases with higher clearance time, except when high VE_C (65%) is assumed – then the direction depends on the endpoint candidate. Overall, the type I error distortion is minimized with the more stringent endpoint definition of consecutive positives over time. The most stringent endpoint ++++ is most robust to varying assumptions on VE_T and the proportion of transient infection. It is sensitive to the clearance time and performs the poorest when the mean clearance time is assumed to be 24 months, which is greater than the post vaccination series follow-up time.

6.2 Power calculations for $VE_P > 0$

Table 1B, Table 2B and Table 3B present the power calculations when 65% vaccine efficacy against chronic infections is assumed. When there is no misclassification error assumed, the endpoint + or ++ typically has the highest power in the presence of VE_T and/or VE_C . An exception is in the setting of high mean clearance time of 24 months and high VE_C of 65%, where the endpoint ++++ has higher power. This is the setting where ++++ showed higher type I error rate than the endpoint + in the previous section. Overall, power decreases as the operational definition becomes more stringent.

The sample size for our study was based on the genotype 1 infection rate of 0.0108/month and the chronic infection proportion of 0.65, which corresponds to the rows in Table 2B where $VE_T = VE_C = 0$ and the proportion of transient infection is 0.35. The power is lower than 80% for all the endpoint candidates, with the highest power given by the endpoint ++++ at 73% in this setting. This likely reflects the imperfect nature of the operational endpoints in capturing the true chronic infections. The assumed cause-specific hazard of chronic infections, $\lambda_{1,2b} = 0.0704$, is not truly reflected in the placebo group using the operational endpoints; in fact, the probability of meeting the endpoint criteria is increased, because there are transient infections that meet the endpoint criteria based on infection statuses. Meeting the endpoint criteria erroneously is least likely to occur with the most stringent definition.

6.3 Effects of misclassification

The last columns of Table 1A–Table 3B present results with misclassification errors. With specificity error, the type I errors decrease. With imperfect HCV tests, it is easier to meet the criteria for the operational definitions when, in fact, the infection is not chronic. For this reason, the imperfect specificity has the most influence on $+$. The power is also most influenced by the misclassification error for the endpoint $+$. This effect is dramatic, for instance, with the power decrease from 95.1% to 28.2% in the setting where genotype 1 infection rate is 0.0108 per month, VE_T is 65%, and the transient infection proportion is 15% (Table 2B), with 5% specificity error.

7. Discussion

In this work, we examined how well the operationally defined HCV infection events can reflect the vaccine effects on chronic infection in a vaccine trial setting. We postulated an underlying multi-state Markov process that includes hypothetical transient and chronic infection states as a tool to assess several infection endpoint candidates. Although a time-homogeneous Markov model may be too simple for the HCV infection-disease process, the purpose was to convey the potential dangers of relying on operationally defined infection endpoints. Even under the simple assumptions of the time-homogeneous Markov model, our results showed some inflation in the type I error in the log-rank test on the vaccine efficacy against chronic infections, in the presence of vaccine efficacy related to transient infections. The worst setting was in the case of high infection rate with high proportion of transient infections, with high vaccine efficacy against transient infections. The type I error rate was up to almost four times the planned rate of 5% (Table 3A). However, the inflation is not as serious as in the HPV setting [18]. The HPV and HCV natural histories are different in that the proportion of chronic infections is high with HCV, compared to HPV where most infections are cleared. For the study power, the imperfect nature of the operational endpoints leads to decreased power. This is likely due to increased event rates in both the vaccine and placebo groups with the operational endpoints based on the serial HCV detection tests, compared to the assumed true chronic infection rates. For instance, the placebo and vaccine group event probabilities for $++...+$ (π_{mp} and π_{mv}) in the setting considered for the sample size determination at the visit months 9, 12, 15, 18, 21, 24 were (0.0211, 0.0209, 0.0204, 0.0204, 0.0206, 0.0213) and (0.00764, 0.00774, 0.00787, 0.00824, 0.00891, 0.0100), respectively, in our model, yielding the power of 0.730. Compare these to the prevalence rates (0.0209, 0.0205, 0.0200, 0.0196, 0.0192, 0.0188) and (0.00736, 0.00731, 0.00726, 0.00720, 0.00715, 0.00710) when the infection rate of 0.00704 is assumed in an exponential distribution. The operational endpoint yields higher event rates.

In considering the endpoints, the advantages of a simple definition for the study endpoint cannot be ignored. The more stringent definition, such as $++...+$, is likely to be more influenced by the loss to follow-up. This endpoint may behave more like the endpoint $+++$ when study non-compliance is considered and requires further consideration as the endpoint candidate.

In this study, only the specificity error was examined in misclassification considerations. The specificity error is probably more serious than the sensitivity error in this evaluation. The specificity error leads to incorrectly identifying an infection status as positive, falsely increasing the number of events that are counted as the chronic infection endpoint, thus leading to higher type I error.

Our work demonstrated the challenges and potential dangers in using an operationally defined infection endpoint in a clinical trial, in the framework of type I and II errors. The properties of these endpoints depended on the natural history of the disease and the vaccine mode of action. For future consideration is an assessment of competing risks. A vaccine that is highly efficacious only against the chronic infections may lead to a slight increase in transient

infections due to competing risks. This was discussed in [27] for the HPV vaccination considerations. Because the majority of HCV infections are chronic, a highly efficacious vaccine against chronic infections may lead to elevation in transient infections, which may compete with chronic infections. A reduction in chronic infection hazard could increase opportunities for transient infections in the vaccinated. It may be helpful to consider the extent of such effect in the HCV setting, when a promising vaccine becomes available.

References

1. Global surveillance and control of hepatitis C. Report of a WHO Consultation organized in collaboration with the Viral Hepatitis Prevention Board, Antwerp, Belgium. *J Viral Hepat* 1999;6(1): 35–47. [PubMed: 10847128]
2. Alter MJ, Kruszon-Moran D, Nainan OV, et al. The prevalence of hepatitis C virus infection in the United States, 1988 through 1994. *N Engl J Med* 1999;341(8):556–562. [PubMed: 10451460]
3. Nainan OV, Alter MJ, Kruszon-Moran D, et al. Hepatitis C virus genotypes and viral concentrations in participants of a general population survey in the United States. *Gastroenterology* 2006;131(2):478–484. [PubMed: 16890602]
4. Kuiken C, Mizokami M, Deleage G, et al. Hepatitis C databases, principles and utility to researchers. *Hepatology* 2006;43(5):1157–1165. [PubMed: 16628639]
5. Flamm SL. Chronic hepatitis C virus infection. *Jama* 2003;289(18):2413–2417. [PubMed: 12746366]
6. Alter MJ, Moyer LA. The importance of preventing hepatitis C virus infection among injection drug users in the United States. *J Acquir Immune Defic Syndr Hum Retrovirol* 1998;18:S6–S10. [PubMed: 9663617]
7. Cox AL, Netski DM, Mosbrugger T, et al. Prospective evaluation of community-acquired acute-phase hepatitis C virus infection. *Clin Infect Dis* 2005;40(7):951–958. [PubMed: 15824985]
8. Amin J, Law MG, Micallef 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(1): 144–150. [PubMed: 16707030]
9. Rustgi VK. The epidemiology of hepatitis C infection in the United States. *J Gastroenterol* 2007;42(7):513–521. [PubMed: 17653645]
10. Garfein RS, Doherty MC, Monterroso ER, et al. Prevalence and incidence of hepatitis C virus infection among young adult injection drug users. *J Acquir Immune Defic Syndr Hum Retrovirol* 1998;18:S11–S19. [PubMed: 9663618]
11. Miller CL, Johnston C, Spittal PM, et al. Opportunities for prevention: hepatitis C prevalence and incidence in a cohort of young injection drug users. *Hepatology* 2002;36(3):737–742. [PubMed: 12198668]
12. Smith RE. Hepatitis C virus therapies. *Nat Rev Drug Discov* 2006;5(9):715–716. [PubMed: 17001802]
13. Chevaliez S, Pawlotsky JM. Hepatitis C virus serologic and virologic tests and clinical diagnosis of HCV-related liver disease. *Int J Med Sci* 2006;3(2):35–40. [PubMed: 16614740]
14. Firbas C, Jilma B, Tauber E, et al. Immunogenicity and safety of a novel therapeutic hepatitis C virus (HCV) peptide vaccine: a randomized, placebo controlled trial for dose optimization in 128 healthy subjects. *Vaccine* 2006;24(20):4343–4353. [PubMed: 16581161]
15. Wintermeyer P, Wands JR. Vaccines to prevent chronic hepatitis C virus infection: current experimental and preclinical developments. *J Gastroenterol* 2007;42(6):424–432. [PubMed: 17671756]
16. Houghton M, Abrignani S. Prospects for a vaccine against the hepatitis C virus. *Nature* 2005;436(7053):961–966. [PubMed: 16107836]
17. Chen SL, Morgan TR. The natural history of hepatitis C virus (HCV) infection. *Int J Med Sci* 2006;3(2):47–52. [PubMed: 16614742]
18. Kang M, Lagakos SW. Evaluation of log-rank tests for infrequent observations from a multi-state process, with application to HPV vaccine efficacy. *Stat Med* 2004;23(23):3681–3696. [PubMed: 15534891]

19. Cox, DR.; Miller, HD. The theory of stochastic processes. London: Chapman and Hall; 1987.
20. Hagan H, Thiede H, Weiss NS, et al. Sharing of drug preparation equipment as a risk factor for hepatitis C. *Am J Public Health* 2001;91(1):42–46. [PubMed: 11189822]
21. Hagan H, Thiede H, Des Jarlais DC. Hepatitis C virus infection among injection drug users: survival analysis of time to seroconversion. *Epidemiology* 2004;15(5):543–549. [PubMed: 15308953]
22. Thorpe LE, Ouellet LJ, Hershow R, et al. Risk of hepatitis C virus infection among young adult injection drug users who share injection equipment. *Am J Epidemiol* 2002;155(7):645–653. [PubMed: 11914192]
23. Kamal SM, Moustafa KN, Chen J, et al. Duration of peginterferon therapy in acute hepatitis C: a randomized trial. *Hepatology* 2006;43(5):923–931. [PubMed: 16628640]
24. Krajden M, Ziermann R, Khan A, et al. Qualitative detection of hepatitis C virus RNA: comparison of analytical sensitivity, clinical performance, and workflow of the Cobas Amplicor HCV test version 2.0 and the HCV RNA transcription-mediated amplification qualitative assay. *J Clin Microbiol* 2002;40(8):2903–2907. [PubMed: 12149349]
25. Pawlotsky JM. Use and interpretation of virological tests for hepatitis C. *Hepatology* 2002;36(5 Suppl 1):S65–S73. [PubMed: 12407578]
26. Freedman LS. Tables of the number of patients required in clinical trials using the logrank test. *Stat Med* 1982;1(2):121–129. [PubMed: 7187087]
27. Kang M, Lagakos SW. Evaluating the role of human papillomavirus vaccine in cervical cancer prevention. *Stat Methods Med Res* 2004;13(2):139–155. [PubMed: 15068258]

Acknowledgments

We are grateful to Bruce Scharschmidt, Olaf Zent and Stephan Weber for fruitful and stimulating discussions. We also thank Michael Hughes for his insightful comments on the manuscript draft.

Minhee Kang is supported by the Statistical and Data Management Center of the AIDS Clinical Trials Group, under the National Institute of Allergy and Infectious Diseases grant No. 1 U01 AI068634.

Figure 1.
Unobservable underlying process.

Table 1

Table 1A: Type I error rates for hazard rate for genotype 1 infection of 0.00702/month									
Null Hypothesis	Proportion of transient infection	Operational Definition	HCV Detection Specificity Error						
			None						
			Clearance Time (Months)						
			6		24		6		24
H_0 ; $VE_P=0$, $VE_I=0$, $VE_C=0$	0.15	+	0.050	0.050	0.050	0.050	0.050	0.050	0.050
		++	0.050	0.050	0.050	0.050	0.050	0.050	0.050
		+++	0.050	0.050	0.050	0.050	0.050	0.050	0.050
		++...+	0.050	0.050	0.050	0.050	0.050	0.050	0.050
H_0 ; $VE_P=0$, $VE_I=25\%$, $VE_C=25\%$	0.15	+	0.052	0.052	0.050	0.050	0.050	0.050	
		++	0.051	0.052	0.051	0.051	0.051	0.052	0.052
		+++	0.051	0.052	0.051	0.051	0.052	0.052	0.051
		++...+	0.050	0.051	0.050	0.050	0.050	0.051	0.051
H_0 ; $VE_P=0$, $VE_I=65\%$, $VE_C=0$	0.15	+	0.060	0.064	0.051	0.051	0.051	0.052	
		++	0.053	0.060	0.053	0.053	0.053	0.059	0.060
		+++	0.051	0.057	0.051	0.051	0.051	0.057	0.057
		++...+	0.050	0.054	0.050	0.050	0.050	0.054	0.054
H_0 ; $VE_P=0$, $VE_I=0$, $VE_C=65\%$	0.15	+	0.052	0.050	0.050	0.050	0.050	0.050	
		++	0.055	0.052	0.054	0.051	0.054	0.051	0.050
		+++	0.052	0.053	0.052	0.052	0.052	0.053	0.053
		++...+	0.050	0.054	0.050	0.050	0.050	0.054	0.054
H_0 ; $VE_P=0$, $VE_I=0$, $VE_C=65\%$	0.35	+	0.065	0.052	0.052	0.052	0.052	0.050	
		++	0.080	0.062	0.073	0.073	0.073	0.059	0.059
		+++	0.065	0.070	0.065	0.070	0.065	0.068	0.068
		++...+	0.053	0.076	0.053	0.076	0.053	0.074	0.074

Table 1B: Power calculations for hazard rate for genotype 1 infection of 0.00702/month

Alternative Hypothesis	Proportion of transient infection	Operational Definition	HCV Detection Specificity Error				
			None		5%		
			Clearance Time (Months)		Clearance Time (Months)		
			6	24	6	24	24
H_A : $VE_P=65\%$, $VE_T=0$, $VE_C=0$	0.15	+	0.730	0.713	0.138		0.137
		++	0.716	0.676	0.611		0.579
		+++	0.679	0.629	0.654		0.606
		++...+	0.699	0.649	0.679		0.628
H_A : $VE_P=65\%$, $VE_T=25\%$, $VE_C=25\%$	0.35	+	0.487	0.454	0.100		0.098
		++	0.516	0.436	0.420		0.364
		+++	0.513	0.409	0.487		0.391
		++...+	0.565	0.447	0.544		0.429
H_A : $VE_P=65\%$, $VE_T=25\%$, $VE_C=25\%$	0.15	+	0.795	0.782	0.151		0.150
		++	0.771	0.745	0.668		0.647
		+++	0.721	0.696	0.698		0.673
		++...+	0.719	0.708	0.699		0.688
H_A : $VE_P=65\%$, $VE_T=65\%$, $VE_C=0$	0.35	+	0.668	0.636	0.124		0.123
		++	0.672	0.612	0.560		0.516
		+++	0.630	0.574	0.605		0.550
		++...+	0.622	0.600	0.602		0.580
H_A : $VE_P=65\%$, $VE_T=65\%$, $VE_C=0$	0.15	+	0.859	0.865	0.167		0.172
		++	0.800	0.819	0.702		0.726
		+++	0.732	0.760	0.712		0.739
		++...+	0.720	0.748	0.701		0.729
H_A : $VE_P=65\%$, $VE_T=65\%$, $VE_C=0$	0.35	+	0.845	0.862	0.159		0.170
		++	0.756	0.807	0.650		0.711
		+++	0.663	0.738	0.644		0.717
		++...+	0.627	0.707	0.609		0.688
H_A : $VE_P=65\%$, $VE_T=0$, $VE_C=65\%$	0.15	+	0.799	0.741	0.152		0.142
		++	0.811	0.745	0.706		0.643
		+++	0.753	0.720	0.730		0.694
		++...+	0.730	0.747	0.712		0.727
H_A : $VE_P=65\%$, $VE_T=0$, $VE_C=65\%$	0.35	+	0.680	0.526	0.126		0.108
		++	0.784	0.610	0.660		0.506
		+++	0.719	0.634	0.695		0.603
		++...+	0.657	0.703	0.639		0.680

Table 2

Table 2A: Type I error rates for hazard rate for genotype 1 infection of 0.0108/month

Null Hypothesis	Proportion of transient infection	Operational Definition	HCV Detection Specificity Error					
			None		5%			
			Clearance Time (Months)		Clearance Time (Months)			
			6	24	6	24	6	24
$H_0: VE_P=0,$ $VE_I=0,$ $VE_C=0$	0.15	+	0.050	0.050	0.050	0.050	0.050	0.050
		++	0.050	0.050	0.050	0.050	0.050	0.050
		+++	0.050	0.050	0.050	0.050	0.050	0.050
		++...+	0.050	0.050	0.050	0.050	0.050	0.050
$H_0: VE_P=0,$ $VE_I=25\%,$ $VE_C=25\%$	0.15	+	0.053	0.053	0.051	0.051	0.051	0.051
		++	0.052	0.053	0.052	0.052	0.052	0.052
		+++	0.051	0.052	0.051	0.051	0.051	0.052
		++...+	0.050	0.052	0.050	0.050	0.050	0.052
$H_0: VE_P=0,$ $VE_I=65\%,$ $VE_C=0$	0.15	+	0.064	0.069	0.053	0.053	0.054	0.054
		++	0.055	0.064	0.054	0.054	0.062	0.062
		+++	0.051	0.060	0.051	0.052	0.059	0.059
		++...+	0.050	0.055	0.050	0.050	0.055	0.055
$H_0: VE_P=0,$ $VE_I=0,$ $VE_C=65\%$	0.35	+	0.139	0.172	0.065	0.065	0.071	0.071
		++	0.082	0.139	0.080	0.080	0.127	0.127
		+++	0.061	0.113	0.061	0.061	0.111	0.111
		++...+	0.051	0.085	0.051	0.051	0.084	0.084
$H_0: VE_P=0,$ $VE_I=0,$ $VE_C=65\%$	0.15	+	0.053	0.050	0.051	0.051	0.050	0.050
		++	0.056	0.053	0.055	0.052	0.052	0.052
		+++	0.053	0.055	0.053	0.053	0.054	0.054
		++...+	0.050	0.056	0.050	0.050	0.055	0.055
$H_0: VE_P=0,$ $VE_I=0,$ $VE_C=65\%$	0.35	+	0.070	0.053	0.054	0.054	0.051	0.051
		++	0.093	0.067	0.084	0.084	0.063	0.063
		+++	0.072	0.078	0.072	0.072	0.075	0.075
		++...+	0.053	0.087	0.054	0.054	0.085	0.085

Table 2B: Power calculations for hazard rate for genotype 1 infection of 0.0108/month

Alternative Hypothesis	Proportion of transient infection	Operational Definition	HCV Detection Specificity Error				
			None		5%		
			Clearance Time (Months)		Clearance Time (Months)		
			6	24	6	24	24
H_A : $VE_P=65\%$, $VE_T=0$, $VE_C=0$	0.15	+	0.863	0.848	0.223	0.219	0.219
		++	0.857	0.823	0.787	0.753	0.753
		+++	0.832	0.785	0.811	0.764	0.764
		++...+	0.850	0.804	0.833	0.785	0.785
H_A : $VE_P=65\%$, $VE_T=25\%$, $VE_C=25\%$	0.35	+	0.626	0.585	0.147	0.143	0.143
		++	0.670	0.572	0.581	0.501	0.501
		+++	0.672	0.546	0.644	0.525	0.525
		++...+	0.730	0.593	0.708	0.572	0.572
H_A : $VE_P=65\%$, $VE_T=25\%$, $VE_C=25\%$	0.15	+	0.912	0.902	0.249	0.247	0.247
		++	0.899	0.879	0.838	0.817	0.817
		+++	0.866	0.845	0.849	0.826	0.826
		++...+	0.865	0.855	0.849	0.839	0.839
H_A : $VE_P=65\%$, $VE_T=65\%$, $VE_C=0$	0.35	+	0.812	0.780	0.197	0.192	0.192
		++	0.824	0.764	0.742	0.687	0.687
		+++	0.792	0.732	0.769	0.709	0.709
		++...+	0.786	0.760	0.767	0.740	0.740
H_A : $VE_P=65\%$, $VE_T=65\%$, $VE_C=0$	0.15	+	0.951	0.954	0.282	0.290	0.290
		++	0.919	0.931	0.867	0.882	0.882
		+++	0.875	0.894	0.860	0.880	0.880
		++...+	0.866	0.886	0.851	0.872	0.872
H_A : $VE_P=65\%$, $VE_T=65\%$, $VE_C=0$	0.35	+	0.944	0.953	0.267	0.286	0.286
		++	0.892	0.923	0.829	0.872	0.872
		+++	0.822	0.879	0.806	0.864	0.864
		++...+	0.790	0.855	0.773	0.840	0.840
H_A : $VE_P=65\%$, $VE_T=0$, $VE_C=65\%$	0.15	+	0.914	0.871	0.251	0.230	0.230
		++	0.926	0.879	0.870	0.813	0.813
		+++	0.890	0.864	0.875	0.845	0.845
		++...+	0.874	0.886	0.859	0.870	0.870
H_A : $VE_P=65\%$, $VE_T=0$, $VE_C=65\%$	0.35	+	0.823	0.667	0.200	0.161	0.161
		++	0.911	0.763	0.838	0.677	0.677
		+++	0.868	0.792	0.851	0.764	0.764
		++...+	0.817	0.854	0.801	0.835	0.835

Table 3

Table 3A: Type I error rates for hazard rate for genotype 1 infection of 0.0149/month

Null Hypothesis	Proportion of transient infection	Operational Definition	HCV Detection Specificity Error					
			None		5%			
			Clearance Time (Months)		Clearance Time (Months)			
			6	24	6	24	6	24
$H_0: VE_P=0,$ $VE_I=0,$ $VE_C=0$	0.15	+	0.050	0.050	0.050	0.050	0.050	0.050
		++	0.050	0.050	0.050	0.050	0.050	0.050
		+++	0.050	0.050	0.050	0.050	0.050	0.050
		++...+	0.050	0.050	0.050	0.050	0.050	0.050
$H_0: VE_P=0,$ $VE_I=25\%,$ $VE_C=25\%$	0.15	+	0.053	0.054	0.051	0.051	0.051	0.051
		++	0.052	0.053	0.052	0.053	0.052	0.053
		+++	0.051	0.053	0.051	0.051	0.051	0.053
		++...+	0.050	0.052	0.050	0.050	0.050	0.052
$H_0: VE_P=0,$ $VE_I=65\%,$ $VE_C=0$	0.15	+	0.066	0.073	0.054	0.054	0.055	0.055
		++	0.056	0.067	0.056	0.065	0.065	0.065
		+++	0.052	0.062	0.052	0.062	0.062	0.062
		++...+	0.050	0.056	0.050	0.056	0.050	0.056
$H_0: VE_P=0,$ $VE_I=0,$ $VE_C=65\%$	0.15	+	0.159	0.198	0.072	0.081	0.072	0.081
		++	0.090	0.160	0.088	0.148	0.088	0.148
		+++	0.063	0.129	0.064	0.126	0.064	0.126
		++...+	0.051	0.093	0.052	0.092	0.052	0.092
$H_0: VE_P=0,$ $VE_I=0,$ $VE_C=65\%$	0.35	+	0.054	0.051	0.051	0.050	0.051	0.050
		++	0.058	0.053	0.056	0.053	0.056	0.053
		+++	0.054	0.056	0.054	0.055	0.054	0.055
		++...+	0.050	0.057	0.051	0.057	0.051	0.057
$H_0: VE_P=0,$ $VE_I=0,$ $VE_C=65\%$	0.35	+	0.075	0.053	0.055	0.051	0.055	0.051
		++	0.104	0.071	0.093	0.066	0.093	0.066
		+++	0.078	0.086	0.078	0.082	0.078	0.082
		++...+	0.054	0.097	0.054	0.094	0.054	0.094

Table 3B: Power calculations for hazard rate for genotype 1 infection of 0.0149/month

Alternative Hypothesis	Proportion of transient infection	Operational Definition	HCV Detection Specificity Error				
			None		5%		
			Clearance Time (Months)		Clearance Time (Months)		
			6	24	6	24	24
H_A : $VE_P=65\%$, $VE_T=0$, $VE_C=0$	0.15	+	0.926	0.915	0.315		0.308
		++	0.927	0.900	0.883		0.852
		+++	0.912	0.875	0.897		0.858
		++...+	0.926	0.890	0.914		0.876
H_A : $VE_P=65\%$, $VE_T=25\%$, $VE_C=25\%$	0.35	+	0.719	0.675	0.200		0.191
		++	0.772	0.669	0.697		0.604
		+++	0.782	0.649	0.756		0.626
		++...+	0.836	0.701	0.817		0.679
H_A : $VE_P=65\%$, $VE_T=25\%$, $VE_C=25\%$	0.15	+	0.959	0.953	0.356		0.350
		++	0.955	0.941	0.921		0.904
		+++	0.936	0.921	0.925		0.907
		++...+	0.936	0.928	0.925		0.917
H_A : $VE_P=65\%$, $VE_T=65\%$, $VE_C=0$	0.35	+	0.889	0.860	0.277		0.267
		++	0.905	0.853	0.850		0.796
		+++	0.885	0.832	0.868		0.811
		++...+	0.881	0.857	0.866		0.840
H_A : $VE_P=65\%$, $VE_T=65\%$, $VE_C=0$	0.15	+	0.982	0.983	0.405		0.416
		++	0.967	0.972	0.940		0.949
		+++	0.942	0.953	0.932		0.945
		++...+	0.936	0.948	0.926		0.939
H_A : $VE_P=65\%$, $VE_T=65\%$, $VE_C=0$	0.35	+	0.979	0.982	0.386		0.410
		++	0.952	0.969	0.917		0.943
		+++	0.907	0.944	0.896		0.935
		++...+	0.884	0.929	0.871		0.918
H_A : $VE_P=65\%$, $VE_T=0$, $VE_C=65\%$	0.15	+	0.961	0.932	0.359		0.325
		++	0.971	0.941	0.942		0.901
		+++	0.952	0.934	0.943		0.921
		++...+	0.942	0.948	0.932		0.939
H_A : $VE_P=0$, $VE_C=65\%$	0.35	+	0.898	0.758	0.283		0.219
		++	0.964	0.852	0.924		0.786
		+++	0.940	0.882	0.929		0.869
		++...+	0.905	0.929	0.893		0.916