

Use of early longitudinal viral load as a surrogate to the virologic endpoint in Hepatitis C: a semi-parametric mixed effect approach using SAS.

Igwebuike Enweonye, B&D Life Sciences – Clinical Research, Brussels, Belgium
Jan Serroyen, Methodology and Statistics, Maastricht University, The Netherlands
Ziv Shkedy, Interuniversity Center for Biostat & Stat Bioinformatics, Hasselt, Belgium
Mikael Le Bouter, B&D Life Sciences – Clinical Research, Brussels, Belgium
Adetayo Kasim, Wolfson Research Institute, Durham University, Durham, UK
Willem Talloen, Janssen Pharmaceutical Company of J&J, Beerse, Belgium
Tony Vangeneugden, Janssen Pharmaceutical Company of J&J, Beerse, Belgium

ABSTRACT

In the management of Hepatitis C Virus (HCV) infection, there is a great interest in early monitoring of viral load since early emergence of resistant strains can jeopardize the treatment outcome. In this paper we present an ongoing research to address the question whether early longitudinal viral loads can be used as a surrogate in discriminating responders / non-responders at the end of the trial. The data analyzed in this paper are from a completed, randomized, phase II, clinical study, exploring the effect of a certain protease inhibitor (PI) therapy in combination with the standard of care on efficacy, safety, tolerability, PK and PK/PD of the therapy. We propose a semi-parametric mixed effect model which is used to model the longitudinal sequence up to week 12 of the subjects. Based on subject specific parameters obtained from the semi-parametric model we estimated the probability to be a responder at the end of the trial. The semi-parametric effect models were implemented with the procedure PROC MIXED in SAS version 9.1.

Key Words: Hepatitis C Virus; Longitudinal viral RNA; Surrogate endpoint; Mixed Effects Models; PROC MIXED in SAS.

1. INTRODUCTION

In the management of Hepatitis C Virus (HCV) infection, there is a great interest in being able to identify early who will have a high benefit or who has a greater probability of failure at later stage. Early monitoring of viral load can guide therapy since early emergence of resistant strains can jeopardize the treatment outcome (J theor Biol. 2010 Sep 10).

Typically in HCV trials, patients receive 24/48 weeks of treatment and are then followed up for 24 weeks before assessing the primary efficacy endpoint. Throughout the duration of the study, HCV viral RNA is continually measured resulting in longitudinal measurements. This paper will present an ongoing research to address the question whether the information obtained from the early longitudinal sequence (week 0 to week 12) of each subject can be used as a surrogate in discriminating the status of the subjects (responders / non-responders for treatment) at the end of the trial. The analysis presented in this paper is based on semi-parametric mixed effect models in which subject-specific longitudinal profiles are fitted using truncated spline models (Ruppert et al. 2003). Such a model approach will allow us to estimate a subject specific rate of change of HCV viral RNA which can be used in order to predict the subject's status at the end of the trial (week 72).

The paper is organized as follows: The data used for the exercise are described in section 2. The semi-parametric mixed effects model used in this paper is discussed in section 3 and software application to the data in section 4. The results are outlined in section 5, and finally, we present discussions and conclusions in section 6.

2. DESCRIPTION OF THE DATA: THE HCV PROTEASE INHIBITOR TRIAL

The data analyzed in this paper are outcome of a completed Phase IIb, randomized, double-blind, placebo-controlled trial to compare the efficacy, tolerability and safety of different regimens with a Direct Antiviral (DA) plus PegIFN α -2a and RBV versus PegIFN α -2a plus RBV alone in adult treatment-naïve subject with genotype 1 HCV infection.

The trial consisted of a screening period of maximum 6 weeks, a 24-week (DA treatment groups) or 48-week (control group) treatment period, followed by a post-treatment follow-up period up to 72 weeks after treatment initiation. Treatment-naïve subjects with chronic genotype 1 HCV infection were included in the trial. Subjects are eligible for the trial when aged between 18 and 70 years old, having a documented chronic genotype 1 HCV infection and a screening plasma HCV RNA level of $> 100,000$ IU/mL among other inclusion criteria described in the study protocol.

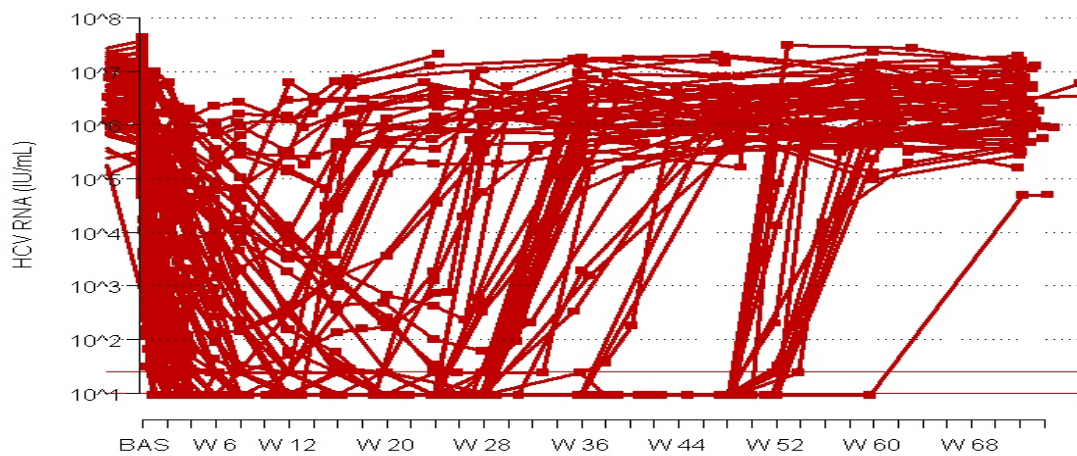
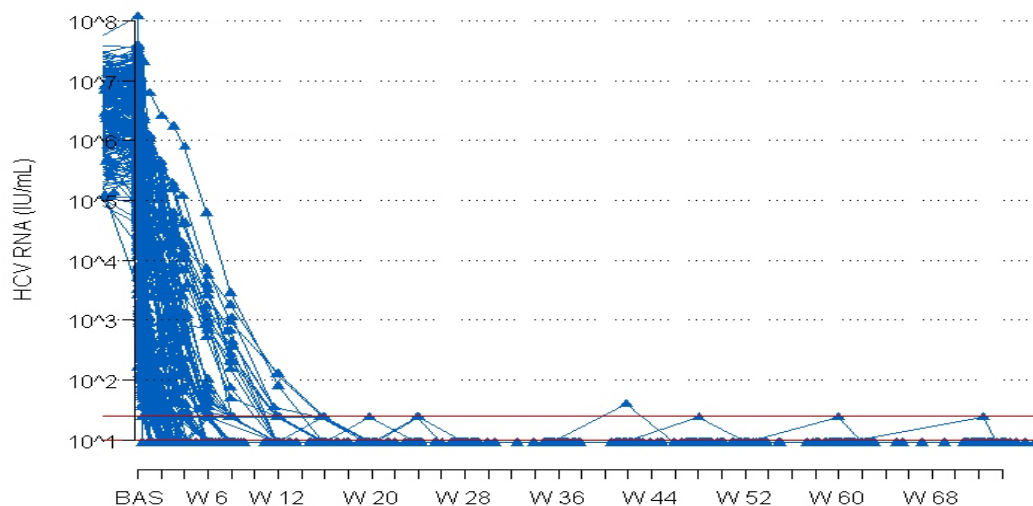


Figure 1 (a): 72-week individual virologic profiles of the subjects who did not achieve sustained virologic response.



(b): 72-week individual virologic profiles of the subjects who achieved sustained virologic response.

HCV RNA was measured during screening and on day 1 before and 4 and 8 hours after dosing. Further measurements were taken on weeks 1, 2, 3, 4, 6, 8, 12, 16, 20, 24, 28, 36, 48, 60 and 72. Subject specific profiles up to 72 weeks are presented in Figure 1 a (non-responders) and b (responders). Note that the semi-parametric mixed effect model is fitted to the data obtained from the first 12 weeks of the trial.

The primary efficacy parameter is the sustained virologic response (SVR) at Week 72. Subjects are defined as responders if they complete the assigned treatment period and achieve undetectable HCV RNA levels at the end of treatment and at Week 72; or prematurely discontinue their assigned treatment period and achieve undetectable HCV RNA levels at the end of treatment and at Week 72. Subjects are considered as failure (non-responders) if they have detectable HCV RNA levels at the end of treatment or detectable or missing HCV RNA levels at Week 72.

We observe (see Figure 1) that the responders begin with rapid decline of wild-type virus followed by a second slower decline until the virus becomes undetectable. In non-responders the virus is not cleared and either rebounds to pretreatment levels, or converges to a lower viral plateau and then returns to pretreatment levels. Most of the non-responders have triphasic or even quadphasic evolution; a situation which poses serious modeling issues. It is evident from the profile that response status can almost be said completely with the naked eyes at treatment week 12 motivating the proposal for the use of surrogate in HCV trials.

3. ESTIMATING THE RATE HCV VIRAL RNA CHANGE USING A SUBJECT SPECIFIC SEMI PARAMETRIC MODELS

3.1. MODEL FORMULATION

Our aim is to develop a flexible subject specific model which allows estimating both subject-specific HCV viral RNA evolution and rate of change over time. Let Y_{ij} be the HCV viral RNA for subject $i (i = 1, 2, \dots, n)$ at time $j (j = 1, \dots, 72)$ in weeks and consider the following model:

$$Y_{ij} = g(t_{ij}) + \varepsilon_{ij} \quad (3.1)$$

Here, $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$ and $E(Y_{ij}) = g(t_{ij})$ is the mean HCV viral RNA level of subject i over time. As we mentioned above, our aim is to estimate $g(t_{ij})$. In this paper we use a truncated spline basis (Ruppert et al. 2003, Maringwa et al. 2008) of the form:

$$g(t_{ij}) = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})t_{ij} + \sum_{k=1}^K b_k (t_{ij} - K_k)_+^p \quad (3.2)$$

Where $k_1, \dots, k_K$ are sets of distinct knots in the range of t_{ij} , with $\mu_+ = \max(0, u)$. The parameters $\beta = [\beta_0 \ \beta_1]'$ are the fixed effects component of the model while b_{0i} and b_{1i} are random specific intercept and slope for which we assume $b_{0i}, b_{1i} \sim N(0, D)$. b_k is a random effect assumed to be normally distributed and independent on b_{0i} and b_{1i} , that is, $b_k \sim N(0, \sigma_b^2)$.

As shown in (Ruppert et al. 2003; Maringwa et al. 2008d) the model specified in (3.2) can be formulated as linear mixed effects model and can be fitted using standard software for mixed models such as procedure MIXED in SAS.

3.2. ESTIMATION OF THE RATE OF CHANGE OF HCV VIRAL RNA LEVELS

The models specified in (3.1) and (3.2) allow us to derive the rate of HCV viral change over time.

Let $g(t_{ij}) = E(Y_{ij} | \beta, b)$ be the mean HCV viral RNA for subject $i (i = 1, 2, \dots, n)$ at time $j (j = 1, \dots, 72)$ in weeks.

The exploratory analysis presented in Section 2 indicates clearly that the HCV viral evolution is not the same for responders and non responders. Therefore a model with a common mean evolution was not considered. In what follows we describe three possible semi-parametric models for HCV viral evolution.

3.2.1 Model 1

The first model that was considered is a model in which the two groups differ in the linear part of the model but they share the same non parametric part. Since our aim is to estimate the first derivative of the mean the truncated power basis ($p=2$) was used. The quadratic power of the truncated basis implies that the smoother will fit peaks and valleys more closely than a linear truncated basis (Ruppert et al. 2003). For this model the mean structure for the HCV viral evolution is given by:

$$\widehat{g}_{ij} = \begin{cases} (\widehat{\beta}_0 + \widehat{b}_{0i}) + (\widehat{\beta}_1 + \widehat{b}_{1i})t_{ij} + \sum_{k=1}^K b_k (t_{ij} - K_k)_+^2, & \text{for NR} \\ (\widehat{\beta}_0 + \widehat{\beta}_{01} + \widehat{b}_{0i}) + (\widehat{\beta}_1 + \widehat{\beta}_{11} + \widehat{b}_{1i})t_{ij} + \sum_{k=1}^K b_k (t_{ij} - K_k)_+^2, & \text{for R} \end{cases} \quad (3.3)$$

Hence, we can derive the rate of HCV viral change as:

$$\widehat{g}_{ij}' = \frac{d}{dt} (\widehat{g}(t_{ij})) = \begin{cases} (\widehat{\beta}_1 + \widehat{b}_{1i}) + \frac{d}{dt} \sum_{k=1}^K b_k (t_{ij} - K_k)_+^2, & \text{for NR} \\ (\widehat{\beta}_1 + \widehat{\beta}_{11} + \widehat{b}_{1i}) + \frac{d}{dt} \sum_{k=1}^K b_k (t_{ij} - K_k)_+^2, & \text{for R} \end{cases} \quad (3.4)$$

3.2.2 Model 2

The second model which was considered is a model in which separate curves are smoothed separately with the same smoothing parameter for the two groups. This can be done by specifying a group specific design matrix for the random effects of the smoother b_k .

$$\widehat{g}_{ij} = \begin{cases} (\widehat{\beta}_0 + \widehat{b}_{0i}) + (\widehat{\beta}_1 + \widehat{b}_{1i})t_{ij} + \sum_{k=1}^K b_k^{NR} (t_{ij} - K_k)_+^2, & \text{for NR} \\ (\widehat{\beta}_0 + \widehat{\beta}_{01} + \widehat{b}_{0i}) + (\widehat{\beta}_1 + \widehat{\beta}_{11} + \widehat{b}_{1i})t_{ij} + \sum_{k=1}^K b_k^R (t_{ij} - K_k)_+^2, & \text{for R} \end{cases} \quad (3.5)$$

Note that compared to model 1 the current model includes two sets of random effects for the smoother b_k^{NR} and b_k^R . This implies that the model uses a group specific smoother in order to estimate the mean response for each group. However, the smoothing parameter is the same for both groups since $var(b_k^{NR}) = var(b_k^R) = \sigma_b^2$ (Maringwa et al. 2008).

The rate of change of viral load $\widehat{g}_{ij}'$ can be derived in a similar way as in (3.4).

3.2.3 Model 3

The third model we considered is a model with subject specific smoothers with the same smoothing parameter. Compare with the previous two models in which subject specific parameters were specified only for the intercept and slope. In the current model the random effects associated with the smoother are subject specific as well. Hence, the mean structure of the model is given by:

$$\widehat{g}_{ij} = \begin{cases} (\widehat{\beta}_0 + \widehat{b}_{0i}) + (\widehat{\beta}_1 + \widehat{b}_{1i})t_{ij} + \sum_{k=1}^K b_{ki} (t_{ij} - K_k)_+^2, & \text{for NR} \\ (\widehat{\beta}_0 + \widehat{\beta}_{01} + \widehat{b}_{0i}) + (\widehat{\beta}_1 + \widehat{\beta}_{11} + \widehat{b}_{1i})t_{ij} + \sum_{k=1}^K b_{ki} (t_{ij} - K_k)_+^2, & \text{for R} \end{cases} \quad (3.6)$$

The rate of change of viral load $\widehat{g}_{ij}'$ can be derived in a similar way as in (3.4). Note that, although, we fitted a subject specific smoother, the total number of parameters in model 3 is the same as in model 1 and 2 since the variance component of the smoother is the same for all subjects.

4. SOFTWARE APPLICATIONS: SEMI PARAMTRIC MIXED EFFECTS MODEL USING PROCEDURE MIXED IN SAS

4.1. SAS CODES AND OUTPUT OF COVARIANCE STRUCTURE FOR MODEL 1

As shown by Ruppert et al. (2003), Maringwa et al. (2008d), the mixed effects model specified in (3.1) – (3.2) can be formulated as a linear mixed model and therefore can be estimated using standard software for linear mixed effects models. The linear mixed effect model formulation of (3.1) – (3.2) are as follow:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b} + \boldsymbol{\varepsilon} \quad (4.1)$$

Where $\mathbf{Y}$ is the m_i -dimensional response vector of measurements for subject $i (i = 1, 2, \dots, n)$, $\mathbf{X}$ and $\mathbf{Z}$ are $m_i \times p$ and $m_i \times q$ dimensional matrices of known covariates respectively, $\boldsymbol{\beta}$ is a p -dimensional vector of fixed effects, $\mathbf{b}$ is q -dimensional subject specific vector of random effects and $\boldsymbol{\varepsilon}$ is an m_i -dimensional vector of residuals. The underlying assumptions for $\mathbf{b}$ and $\boldsymbol{\varepsilon}$ are as previously described in section 3.1.

We use SAS procedure MIXED to estimate the proposed model. The mean structure for the fixed effect component of the model can be specified by:

```
MODEL VLLOGRES= TIME SVR TSCALE*SVR/ NOINT SOLUTION
```

The model has two random components related to the random effects of the model. The first random component accounts for subject heterogeneity (subject specific intercept and slope).

```
random intercept Time / subject= usubjid ;
```

The second random component is specified for the smoother. In this component the design matrix for the smoother is specified. In the code below Z1-Z13 are the columns of the design matrix Z in section 3.1.

```
random z1-z13 / type=toep(1) solution;
```

PhUSE 2011 - SP04

Note that the option “`type=toep(1)`” specifies the covariance matrix of the random effect for the smoother b_K which has a K x K diagonal Toeplitz structure of the form $\sigma_b^2 I_{K \times K}$. The complete code is given below:

```
proc mixed data=datasw12 method=ml order=data asycov covtest;
class svr usubjid ;
model vllogres= Time svr Time*svr/ noint solution
outp=predspw12(rename=(pred=vlspw12)) ;
random z1-z13 / type=toep(1) solution;
random intercept Time / subject= usubjid ;
Ods output covparms =cspsw12;
Ods output solutionR =randefsw12;
Ods output solutionF =fixefsw12;
run;
```

The variance components of the smoother of the model are shown in the panel below.

SAS output for Covariance Parameters: Model 1		
Cov Parm	Subject	Estimate
Variance		91.9032
Intercept	USUBJID	0.4367
Time	USUBJID	0.0534
Residual		0.4407

4.2. SAS CODES AND OUTPUT OF COVARIANCE STRUCTURE FOR MODEL 2

In model 1, the smoother was common for both responders and non-responders. Model 2, includes group specific smoother. This can be done by changing the covariance matrix of the random effects for the smoother which now becomes a block diagonal matrix. This implies that two sets of random effects for the smoother will be estimated: one for the responders and the other for the non-responders.

```
random z1-z13 / type=toep(1) subject=SVR solution;
```

The option “`subject=SVR`” implies that two sets of random effects, one set for responders and the other for non responders will be fitted. However, the same variance component is used for the two sets of random effects. The complete code is given below:

```
proc mixed data=datasw12 method=ml order=data asycov covtest;
class svr usubjid ;
model vllogres= Time svr Time*svr/ noint solution
outp=predspw12(rename=(pred=vlspw12)) ;
random z1-z13 / type=toep(1) subject=SVR solution;
random intercept Time / subject=usubjid ;
Ods output covparms =cspsw12;
Ods output solutionR =randefsw12;
Ods output solutionF =fixefsw12;
run;
```

The variance components of the smoother of the model are shown in the panel below

SAS output for Covariance Parameters: Model 2		
Cov Parm	Subject	Estimate
Variance	SVR	8.8865
Intercept	USUBJID	0.4396
Time	USUBJID	0.0538
Residual		0.4320

4.3. SAS CODES AND OUTPUT OF COVARIANCE STRUCTURE FOR MODEL 3

For model 3, we specify subject specific smoother. This can be done by changing the covariance matrix of the random effects for the smoother.

```
random z1-z13 / type=toep(1) subject=usubjid solution;
```

The option “`subject=usubjid`” implies that a subject specific model will be fitted for each unique subject. Similar to models 1 and 2 the same variance component is used for all random effects associated with the smoother.

The complete code is given below:

```
proc mixed data=datasw12 method=ml order=data asycov covtest;
class svr usubjid ;
model vllogres= Time svr Time*svr/ noint solution
outp=predspw12(rename=(pred=vlspw12)) ;
random z1-z13 / type=toep(1) subject=usubjid solution;
random intercept Time / subject=usubjid ;
Ods output covparms =cspw12;
Ods output solutionR =randefsw12;
Ods output solutionF =fixefspw12;
run;
```

As we mentioned above, model 3 estimates a subject specific smoother. Hence the variance component associated with the smoother has for this model the same unit as the subject.

SAS output for Covariance Parameters: Model 3

Cov Parm	Subject	Estimate
Variance	USUBJID	0.2123
Intercept	USUBJID	0.3409
Time	USUBJID	0.3487
Residual		0.2306

4.4. CONSTRUCTION OF THE DESIGN MATRICES FOR THE RANDOM EFFECTS FOR THE SMOOTHER

The design matrix for the smoother can be constructed using the macros below. For a detailed description of these macros we refer to Ruppert et al. (2003).

```
/*The following macro computes the default set of knots from a vector x of time */
%default_knots(librefknots=work,data=vlprofile,varknots=vldy,numknots=10);

/*Macro uses PROC MIXED to perform a penalized linear spline regression with REML
estimation of the amount of smoothing. */

/*definition basis*/
%default_basis(libref=work,data=vlprofile,x=vldy,y=vllogres,id=usubjid,timehrs=vld
y,grp=analtptn,knotdata=work.knots,knots=knots,power=2);
```

5. RESULTS

5.1. EMPIRICAL AND FITTED DATA

Figures 2 and 3 show the observed data (blue lines) and fitted subject specific models using semi-parametric mixed effects model (brown lines). Note that only data until treatment week 12 was used for the analysis. We notice that the HCV RNA values for responders dropped rather rapidly and stay below the threshold as expected while the HCV RNA values for non-responders are higher. Although for some of the subjects they cross the threshold value (see upper panels in Figure 3).

Based on the subject specific profiles we can derive the rate of change in HCV RNA over time and for any given time point. Our aim is to use the model based prediction for the rate of decline of HCV RNA to predict the response status at week 72.

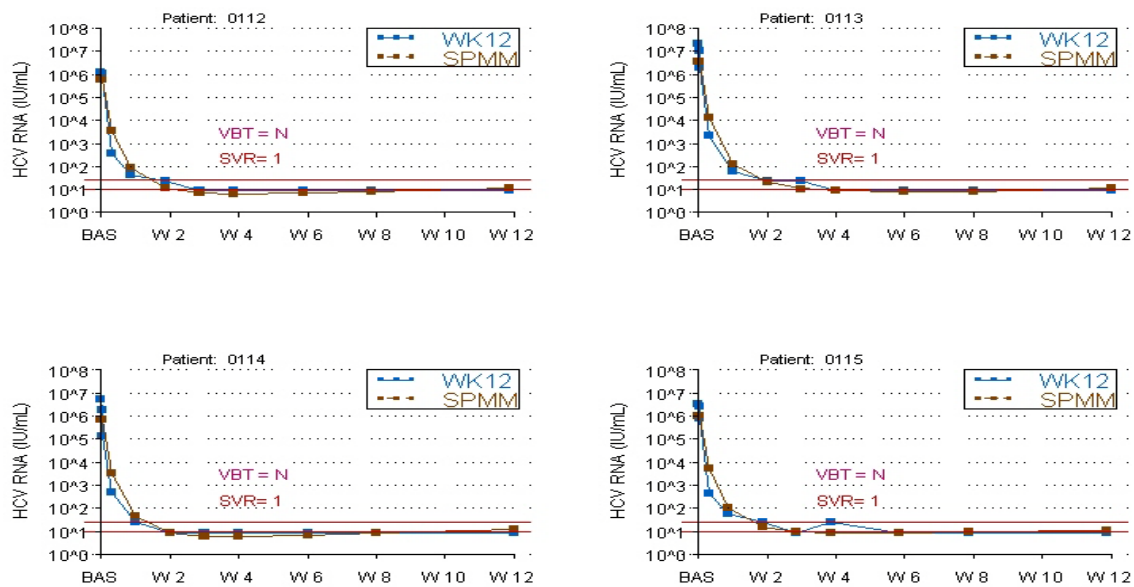


Figure 2: Observed and fitted models for subjects with sustained virologic response (SVR=1) at week 72 and viral break through (VBT (Y/N)) during treatment are indicated. Blue line: observed data. Brown line: subject specific profile obtained from the semi-parametric model.

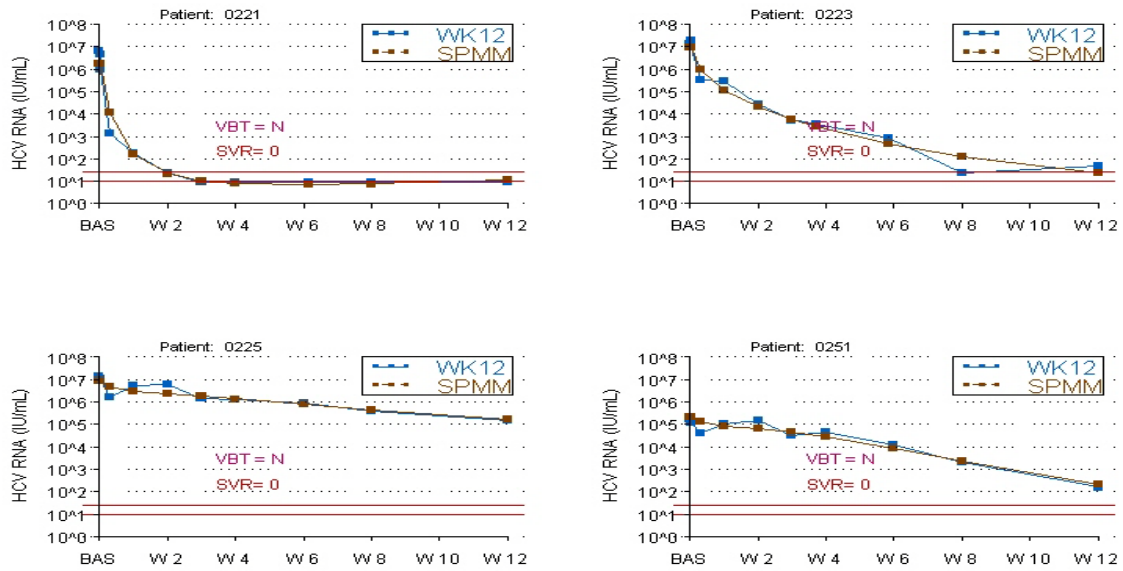


Figure 3: Observed and fitted models for subjects with sustained virologic response (SVR=0) at week 72 and viral break through (VBT (Y/N)) during treatment are indicated. Blue line: observed data. Brown line: subject specific profile obtained from the semi-parametric model.

5.2. INFERENCE FOR THE RATE OF VIRAL CHANGE AND RANDOM EFFECTS

Figure 4 shows the individual rate of viral change at week 12 for two subjects. Further, Figure 5 shows box plots of the rates of viral change at different weeks comparing the responders and non-responders.

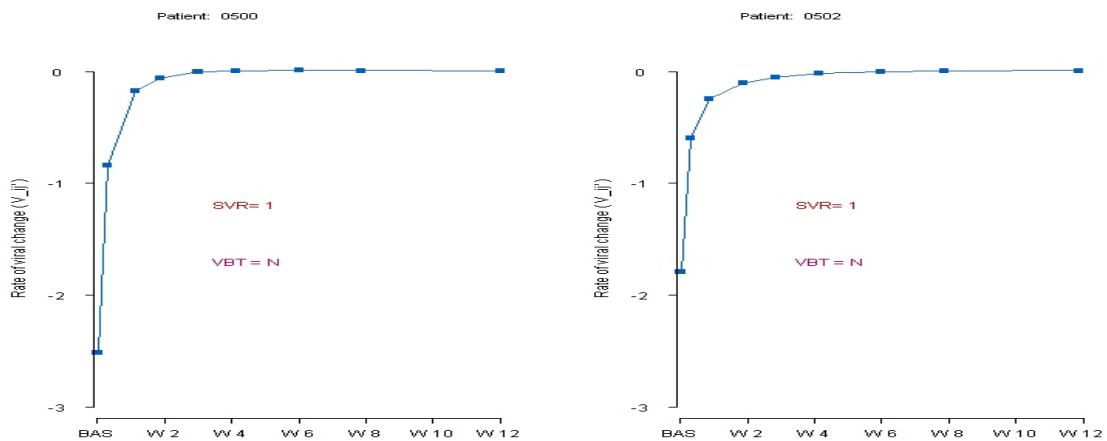


Figure 4: The individual rate of viral change at week 12 for two subjects (0500 and 0502).

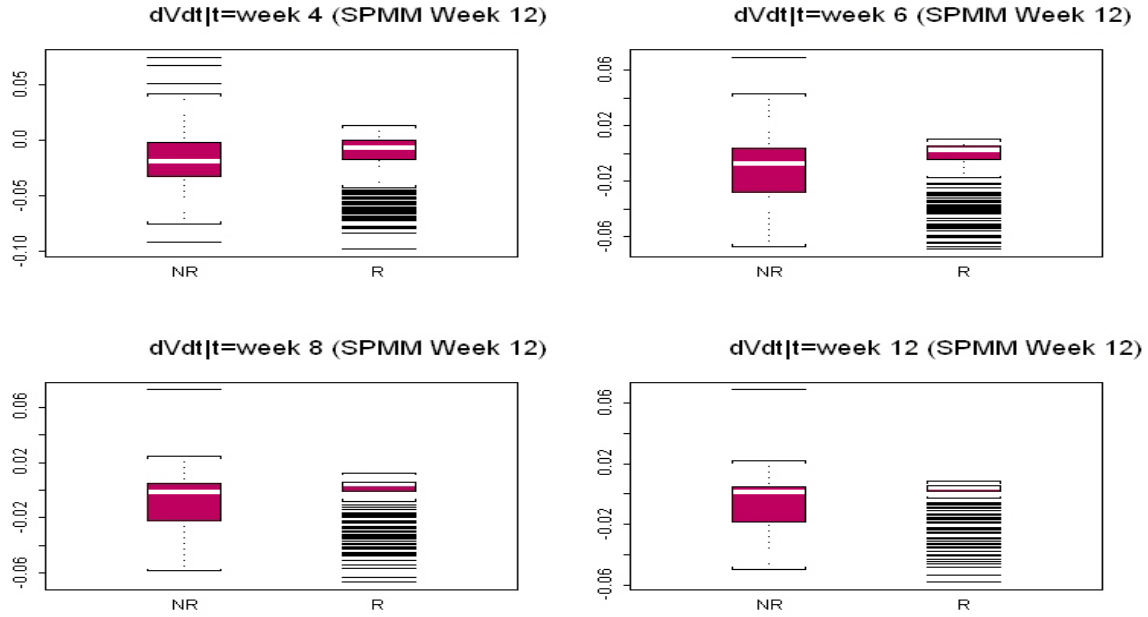


Figure 5: Box plots of the rates of viral change based on data from treatment week 12 and evaluated at different time points. Comparisons are made between responders (R) and non-responders (NR).

Comparing the rate of viral change between responders and non-responders, Figure 5, indicates difference in the two groups. Observe that this effect is reduced for the responders and most notably in week 12. This corresponds to what was earlier observed in Figure 1: individual virologic profiles.

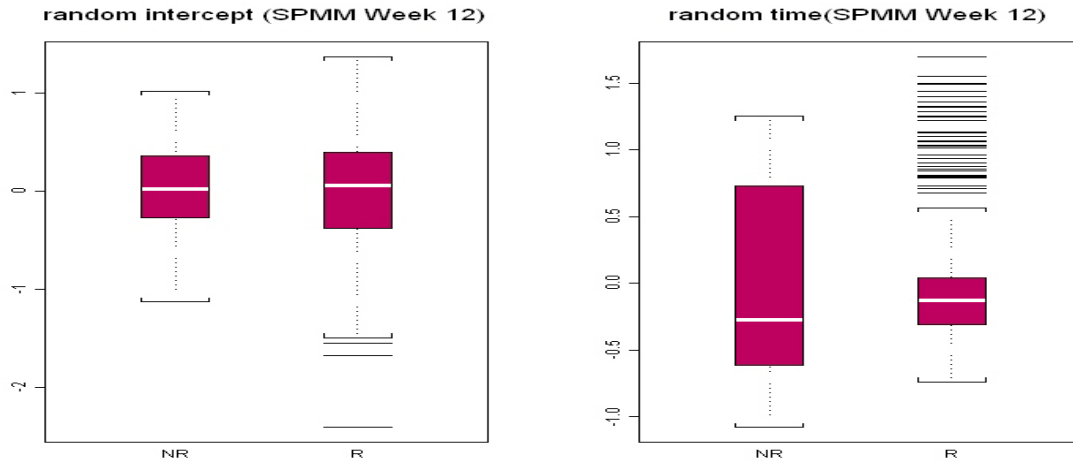


Figure 6: The empirical Bayes estimates for the random intercept and the time slope.

The empirical Bayes estimates for the random intercept and the time slope are shown above. For the random effects, responders (R) and non-responders (NR) are placed side by side. Observe the huge variability among the non-responders for the random time slope meaning that the rate of viral change is very heterogeneous among them.

5.3. PREDICTION OF SUSTAINED VIROLOGIC RESPONSE 24 WEEKS AFTER TREATMENT - SVR24 (OR WEEK 72)

In this section we formulate a predictive model for the status of the subject in the end of the trial at week 72. Let R_i be an indicator variable such that:

$$R_i = \begin{cases} \text{responder (R)}, & \text{if sustained virologic response (SVR = 1)} \\ \text{non-responder (NR)}, & \text{if sustained virologic response (SVR = 0)} \end{cases}$$

Our aim is to model the probability to be a responder as a function of the subject specific parameters estimated from the semi-parametric mixed model discussed in section 3. Let $P[R_i = 1 | \hat{g}'(t=k), \widehat{b_{0i}}]$ be the probability of a subject to be a responder given the viral characteristics – the subject specific rate of change of viral load at time t and the random intercept. We further assume that:

$$g(P[R_i = 1 | \hat{g}'(t=k), \widehat{b_{0i}}]) = \alpha + \beta x \quad (5.1)$$

$g(\cdot)$ is a logit function.

We fitted logistic regression models for sustained virologic response (SVR) as a function of the rate of viral change and random intercept considering 4 different time points at:- week 4, 6, 8 and 12. For each time point the subject specific rate of change is calculated by differentiating (3.6). Once the rate of viral change is calculated it is used together with the random intercept as covariates in the logistic model formulated in (5.1). Subjects are classified as responders or non-responders according to the predicted probability to be a responder obtained from (5.1).

The Receiver Operating Characteristics (ROC) curves (see Figure 7) based on the viral information at week 4, shows 80% sensitivity and 70% specificity. It is evident that the more longitudinal data we use the better the prediction. At week 8 and 12 we achieve a prediction with over 90% sensitivity and near 80% specificity. This is expected since a combination therapy of direct antiviral agent and standard of care causes a lot of viral activities provided resistant strains are not developed. Also figure 1 of individual profiles shows that at week 12 even a naked eye could pick the cases of responders and non-responders.

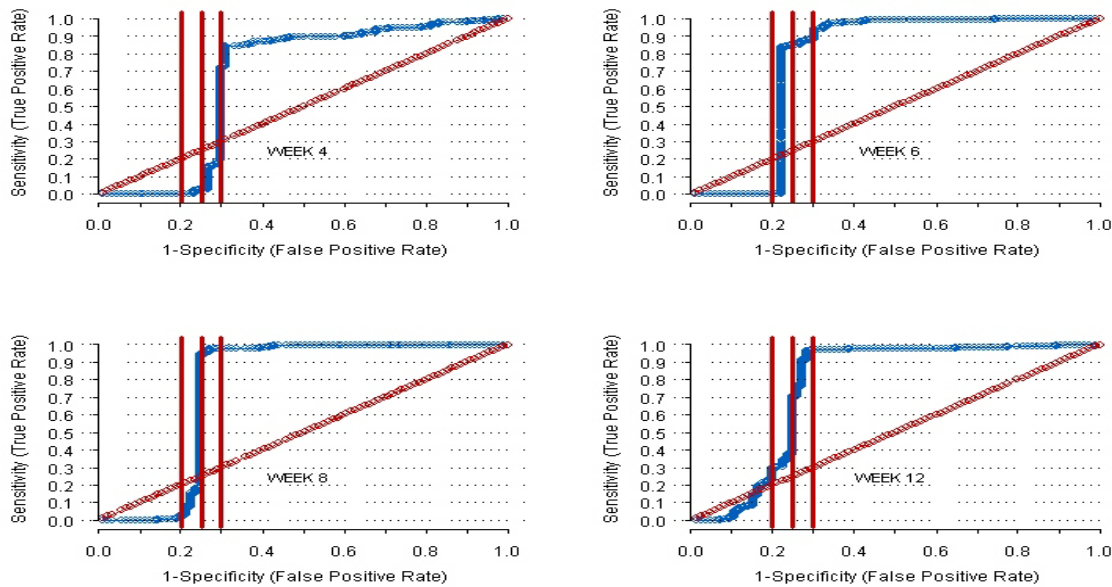


Figure 7: Receiver Operating Characteristics (ROC) curves constructed from fitted logistic model for sustained virologic response (SVR (1/0)) as a function of random intercepts and the rate of viral change at several weeks.

6. DISCUSSIONS AND CONCLUSIONS

This era of management of hepatitis C virus infection using a triple combination of Direct Antiviral (DA), PegIFN and RBV is characterized by rapid wild-type viral decline but also the tendency of the subjects to develop resistant mutations which can jeopardize further treatment. In order to avoid the emergence of resistant strains, there is an interest in being able to identify early subjects who will respond to treatment and those not.

Typically in HCV trials, patients receive 24/48 weeks of treatment and are then followed up for 24 weeks before assessing the primary efficacy endpoint; resulting in longitudinal data. The relevant question here is whether the information obtained from the early longitudinal sequence can be used a surrogate to predict responders and non-responders at the end of the study.

To answer the above question, we used clinical data obtained from a completed Phase IIb, randomized, double-blind, placebo-controlled trial to compare the efficacy, tolerability and safety of different regimens with a Direct Antiviral (DA), plus PegIFN α -2a and RBV versus PegIFN α -2a plus RBV alone in adult treatment-naïve subject with genotype 1 HCV infection. The evolution of viral profiles over time was studied via semi-parametric mixed effects models.

It is observed that sustained virologic response at week 24 (SVR24) post treatment is strongly correlated with early viral activities of the patients at treatment week 12. The model for the prediction of SVR based on the viral information at week 12 is over 90% sensitive and near 80% specific. We conclude that individual level surrogacy based on week 12 viral profile to predict the outcome of interest is possible. However, extrapolation of this information to a new study is a subject of further research among the clinical team.

REFERENCES

- Dahari, H. et al. (2007) Modeling Hepatitis C virus dynamics: Liver regeneration and critical drug efficacy. *Journal of Theoretical Biology*, **247**, 371 – 381.
- Guedj, J. et al. (2010) A perspective on modeling hepatitis C virus infection. *Journal of Viral Hepatitis*, **17**, 825-833.
- Maringwa, J.T. et al. (2008d) Application of semiparametric mixed models and simultaneous confidence bands in a Cardiovascular safety experiment with longitudinal data. *Journal of Biopharmaceutical Statistics*, **18**, 000-000.
- Royston, P., Altman, D.G. (1994) Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling. *Applied Statistics*, **43**, 429-467.
- Ruppert, D. (2002) Selecting the number of knots for penalized splines. *Journal of Computational and Graphical Statistics*, **11**, 735-757.
- Ruppert, D., Carroll, R.J. (2000) Spatially adaptive penalties for spline fitting. *Australian and New Zealand Journal of Statistics*, **42**, 205-223.
- Ruppert, D., Wand, M.P., Carroll, R.J. (2003) *Semiparametric Regression*. Cambridge University Press.
- Verbeke, G., Molenberghs, G. (2000) *Linear Mixed Models for Longitudinal data*. New York: Springer.
- Verbeke, G., Molenberghs, G. (2003) The use of score tests for inference on variance components. *Biometrics*, **59**, 254-262.

ACKNOWLEDGMENTS

The authors are grateful to Janssen Pharmaceutical Companies of Johnson and Johnson Beerse; B&D Life Sciences Brussels; Jerry Welkenhuysen-Gybel; and Adaobi Enweonye, for their support.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Igwebuike Enweonye, Business & Decision Life Sciences Benelux

Sint Lambertusstraat 141, Brussels 1200

Work Phone: +322 774 11 00

Email: igwebuike.enweonye@businessdecision.com

Web: www.businessdecision-lifesciences.com