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

- The use of surrogate endpoints in clinical trials is increasing, necessitating the development of sound statistical methods in the validation process (Burzykowski et al. 2005).

- Although the idea of surrogacy was developed in the context of single trial (Prentice 1989, Freedman et al. 1992, Buyse and Molenberghs 1998), the meta-analytic approach is now a well accepted one.

- The meta-analytic approach allows both individual and trial level surrogacy.

- For our purposes, we shall keep to individual level surrogacy.
Introduction (contd.)

• Important implications of statistical validation of surrogate:
  – A prediction but also discriminative model is needed.
  – Validity of a surrogate = quality of prediction.
  – Model extrapolation to a new treatment (mechanism)?

• 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 evaluated resulting in longitudinal measurements.
Research Question:

- Can the first few Measurements in the Longitudinal Sequence be used as a Surrogate to the Virologic Endpoint in Hepatitis C Virus (HCV) Infection Trials?
The HCV protease inhibitor data description

• Phase IIb, randomized, double-blind, placebo-controlled trial to compare the efficacy, tolerability and safety of different regimens with Direct Antiviral (DA) plus PegIFN-2a and RBV versus PegIFN-2a plus RBV alone in adult treatment-naive subjects with genotype 1 HCV infection.

• Measurements taken at week: 1, 2, 3, 4, 6, 8, 12, 16, 20, 24, 28, 36, 48, 60 and 72.

• Analysis limited to the data obtained from the first 12 weeks.

• The primary efficacy parameter is the sustained virologic response (SVR):
  – SVR is undetectable HCV RNA levels at the end of treatment and at Week 72
The HCV protease inhibitor trial: Non-responders
The HCV protease inhibitor trial: Responders
Individual profiles (Week 72)

Patient Nr.: 0426
Randomisation group 175/12W - NSSB: 1B
Trial Termination: completed

Patient Nr.: 0206
Randomisation group 175/12W - NSSB: 1B
Trial Termination: completed

Patient Nr.: 0371
Randomisation group 175/12W - NSSB: 1A
Trial Termination: discontinued

Patient Nr.: 0001
Randomisation group 175/12W - NSSB: 1A
Trial Termination: completed
Methods: The Semi-Parametric Mixed Effects Model

- Define the viral RNA for patient $i$ at time $t_{ij}$ as:
  \[ y_{ij} = g(t_{ij}) + \varepsilon_{ij} \]
- \( g(t_{ij}) = (\beta_0 + b_{oi}) + (\beta_1 + b_{1i})t_{ij} + \sum_{k=1}^{K} b_k(t_{ij} - k_k)^{\phi} \)
- \( k_1, \ldots, k_k \) are a set of distinct knots in the range of \( t_{ij} \), with \( \mu_+ = \max(0, \mu) \)
- The random specific intercept and slope \( b_{oi}, b_{1i} : N(0, \sigma) \), \( b_k \) is random effect associated with the smoother and assumed normally distributed and independent of \( b_{oi}, b_{1i} \), that is, \( b_k : N(0, \sigma_b^2) \), give SPMM with \( \varepsilon : N(0, \sigma_\varepsilon^2) \).
- \( Y = \left[ y_{ij} \right]_{1 \leq i \leq n, 1 \leq j \leq m} \) (Durban et al., 2005) be the vector of stacked patient specific viral loads
Methods: The Semi-Parametric Mixed Effects Model

- \( X = [1_t_{ij}]_{1 \leq s \leq n, 1 \leq j \leq m_i} \) the corresponding design matrix.
- \( \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \) the vector of fixed effects to be estimated.
- \( Z_i = \left[ (t_{ij} - k_k) \right]^p_{1 \leq j \leq m_i, 1 \leq k \leq K} \) is design matrix for the smoother.
- Analogous to the linear mixed model (Ruppert et al. 2003; Maringwa et al. 2008d):
  \[ Y = X\beta + Zb + \varepsilon \]
- As a consequence, the SPMM can be fitted using standard software for mixed models such as procedure MIXED in SAS®.
Model (1.1) is one in which the two groups differ in the linear part of the model but they share the same non-parametric part.

\[
g_{ij} = \begin{cases} 
(\hat{\beta}_0 + \hat{b}_{0i}) + (\hat{\beta}_1 + \hat{b}_{11})t_{ij} + \sum_{k=1}^{K} b_k (t_{ij} - K_k)^2_+ , & \text{if NR} \\
(\hat{\beta}_0 + \hat{\beta}_{01} + \hat{b}_{11}) + (\hat{\beta}_1 + \hat{\beta}_{11} + \hat{b}_{1i})t_{ij} + \sum_{k=1}^{K} b_k (t_{ij} - K_2)^2_+ , & \text{if R} 
\end{cases}
\]

The truncated basis (\( \varphi = 2 \)) implies that the smoother will fit peaks and valleys more closely than a linear truncated basis (Ruppert et al. 2003).
Methods: The Semi-Parametric Mixed Effects Model

The rate of viral decline becomes:

\[ g_{ij} = \begin{cases} 
(\hat{\beta}_0 + \hat{b}_{0i}) + \frac{d}{dt} \sum_{k=1}^{K} b_k (t_{ij} - k_k)^2, & \text{for NR} \\
(\hat{\beta}_1 + \hat{\beta}_{11} + \hat{b}_{1i}) \frac{d}{dt} \sum_{k=1}^{K} b_k (t_{ij} - k_k)^2, & \text{for R} 
\end{cases} \]

Application of (1.1) in SAS®

The mean structure for the fixed effects component of the model can be specified by:

```sas
MODEL VLLOGRES = Time SVR Time*SVR/ NOINT SOLUTION
```
The model has two random components. The first random component accounts for subject heterogeneity.

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

The second random component specifies the design matrix for the smoother. Z1-Z13 below are the columns of the design matrix Z.

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

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

```
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 =cpspw12;
Ods output solutionR =randefsw12;
Ods output solutionF =fixefspw12;
run;
```
Methods: The Semi-Parametric Mixed Effects Model

- Model (1.2) separate curves smoothed separately with the same smoothing parameter for the two groups. This is done by specifying a group specific design matrix for the random effects of the smoother $\hat{b}_k$.

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

- $\text{var}(\hat{b}_k^{NR}) = \text{var}(\hat{b}_k^{R}) = \sigma_b^2$

- Again, the rate of viral decline can be obtained as in (1.1)
Methods: The Semi-Parametric Mixed Effects Model

Application of (1.2) in SAS®

- Includes group specific smoother by changing the covariance matrix of the random effects for the smoother. This implies that two sets of random effects for the smoother will be estimated.

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

- Model (1.3): model with subject specific smoothers but with the same smoothing parameter.

\[
\hat{g}_{ij} = \begin{cases} 
(\hat{\beta}_0 + \hat{b}_{0i}) + (\hat{\beta}_1 + \hat{b}_{1i})t_{ij} + \sum_{k=1}^{K} \hat{b}_{ki}(t_{ij} - k_k)^2 , & \text{for NR} \\
(\hat{\beta}_0 + \hat{\beta}_{01} + \hat{b}_{0i}) + (\hat{\beta} + \hat{\beta}_{11} + \hat{b}_{1i})t_{ij} + \sum_{k=1}^{K} \hat{b}_{ki}(t_{ij} - k_k)^2 , & \text{for R}
\end{cases}
\]

The rate of decline is obtained in a similar way as in (1.1)
Results: Observed and Fitted data (SVR= 1 )
Results: Observed and Fitted data (SVR= 0 )
Results: Inference for the random effects
Results: Inference for the rate of viral decline

$dV/dt|_{t=\text{week 4 (SPMM Week 12)}}$

$dV/dt|_{t=\text{week 6 (SPMM Week 12)}}$

$dV/dt|_{t=\text{week 8 (SPMM Week 12)}}$

$dV/dt|_{t=\text{week 12 (SPMM Week 12)}}$
Prediction of SVR using the logistic regression

- Let
  \[ R_i = \begin{cases} 
  \text{responder(R)}, & \text{if SVR}=1 \\
  \text{non-responder(NR)}, & \text{if SVR}=0 
  \end{cases} \]

- Aim is to model the probability to be a responder as a function of the subject specific parameters.

- \[ p = P \left[ R_i = 1 \bigg| \hat{g}'(t = k), \hat{b}_o \right] \text{ given the viral characteristics.} \]

- Using logit link, \[ \log \left( \frac{p}{1 - p} \right) = \alpha + \beta x; \text{ and, } p = \frac{\exp (\alpha + \beta x)}{1 + \exp (\alpha + \beta x)} \]

- We fitted logistic regression models for SVR as a function of the rate of viral change and random intercept at different time points
Results: Prediction of sustained virologic response
Conclusions

- This era of HCV treatment using a triple therapy of Direct Antiviral (DA), PegIFN and RBV leads to rapid wild-type viral decline but also to development of resistant mutations.

- 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 evaluated resulting in longitudinal measurements.

- The evolution of viral profiles over time was studied via semi-parametric mixed effects models.
Conclusions

• 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, surrogate extrapolation to a new treatment (mechanism) is a subject of further research among the clinical team.
Acknowledgements

- Paul Meyvisch (J&J Director)
- Ziv Shkedey (Promoter)
- Willem Talloen (collaborator)
- Adetayo Kasim (collaborator)
- Jan Serroyen (collaborator)
- Mikael Le Bouter (collaborator)
- Adaobi Enweonye (Technical Support)
- Tony Vangeneugden (J&J Director and collaborator)
- Jerry Welkenhuysen-Gybel (B&D LS Managing Director)