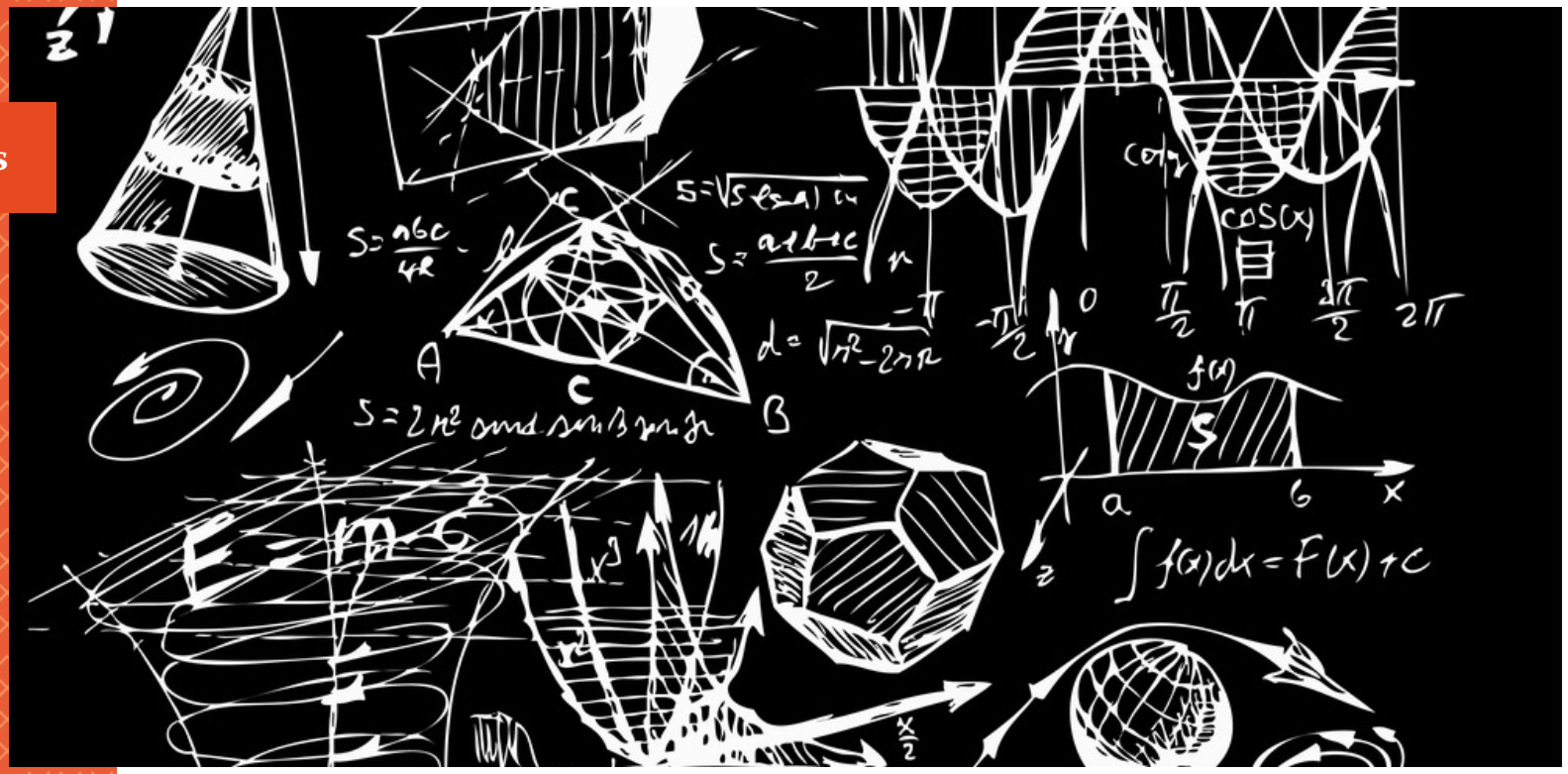
Novartis



Set Control limit at Release for Stability Risk

Xiang-Yi (Gregory) Chen
Global AS&T/QC/OpEx, NTO Quality

2018 NCS Conference, Paris

NOVARTIS

Acknowledge

Prof. A. W. van der Vaart

Preface

The approaches discussed here are not the one that currently implemented in Novartis.

Question of Interests

How to do set *a control limit at time zero* (release) to provide assurance for *the concerned stability risk*? For brevity, call this interim or internal release control limit as IRL.

Typical Motivating Scenario:

When release and stability spec are the same, and a trend is expected towards the direction with regulation, stringency of that spec at release and at the end of shelf life (EOSL) can be highly unbalanced.

Usual Objective

Concerned stability risk => ***Low passing rate at EOSL due to trend***

Particularly, we want to exam what is the "honest" statistical solution to the desired prescribed probabilistic statement, and its practicality.

Type of Stability Risk

A future batch passing the determined IRL would ideally have

- *High passing rate at EOSL*
- *High joint passing rate across the following up stability study (at a pre-defined time grid)*
- *High confidence that the batch mean at EOSL is within stability spec*
- … …

Naturally, different objectives leading to different solution.

- Is the formulated objective/question acceptable and relevant?
- Is the proposed solution an honest answer to its probabilistic prescription?
- Is the resulting IRL implementable?

Benchmark Solutions

Allen, Dukes and Gerger (ADG) 1991; G.C.Wei 1998

*Only consider degradation and lower specification for illustration

ADG (1991)

$$IRL = LSS - \hat{b}T + t_{0.95,n^*}\sqrt{T^2 s_{\hat{b}}^2 + \frac{\hat{\sigma}_e^2}{k}}, \quad (1)$$

where LSS refers to lower stability specification, T is the shelf life (unit in month), $\hat{b}$ is an estimate of overall slope across batches in the data, $s_{\hat{b}}$ is the standard error of estimate $\hat{b}$, $\hat{\sigma}_e^2$ is the assay variance, k is the number of replicates to calculate the average as a reportable value (to be compared with RS or SS), and $t_{0.95,n^*}$ is the 95% percentile of a student-t distribution with n^* degrees of freedom. n^* is calculated by the Satterthwaite approximation. We should note that $\hat{\sigma}_e^2$ is not the residual variance from a simple regression model where $\hat{b}$ and $s_{\hat{b}}$ are derived.

*Overarching assumption is that the stability behaviour of a future batch would follow the overall trend in the current data. **The objectives** is to find the minimum batch mean at release such that its batch mean at time T would not be significantly <= LSS at 0.05 level,*

Drawback of ADG approach

- The objective looks for a batch mean which can never be observed in reality, while determined IRL would be applied to individuals.

- Only overall trend is concerned here, the batch-to-batch variation of the stability behavior is overlooked.

Observation Model

Standard Random-Coefficients Model

$$Y_{it} = a + \alpha_i + (b + \beta_i)t + \varepsilon_{it}. \tag{2}$$

Under a regular setup, we assume the overall coefficients a and b are fixed, batch random intercept $\alpha_i \sim N(0, \sigma_\alpha^2)$, random slope $\beta_i \sim N(0, \sigma_\beta^2)$, residual error $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$, and α_i, β_i and ε_k are mutually independent for $\forall i, j, k$.

For any given t_1 and t_2, and let $t_1 < t_2$

$$\begin{pmatrix} Y_{it_1} \\ Y_{it_2} \end{pmatrix} \sim N\left(\begin{pmatrix} a + bt_1 \\ a + bt_2 \end{pmatrix}, \begin{pmatrix} \sigma_\alpha^2 + \sigma_\beta^2 t_1^2 + \sigma_\varepsilon^2 & \sigma_\alpha^2 + \sigma_\beta^2 t_1 t_2 \\ \sigma_\alpha^2 + \sigma_\beta^2 t_1 t_2 & \sigma_\alpha^2 + \sigma_\beta^2 t_2^2 + \sigma_\varepsilon^2 \end{pmatrix} \right) \tag{3}$$

Most Desired Question: CP Formulation

Assure high passing at EOSL of a batch based on its release measure, similar to one objectives of Wei (1998). $LSS = \gamma$, $IRL = \eta$

$$P_\theta(Y_{it_2} \geqslant \gamma | Y_{it_1} \geqslant \eta) = q \tag{4}$$

where $\theta \subseteq \{a, b, \sigma_\alpha, \sigma_\beta, \sigma_\varepsilon\}$, and q is the desired level of assurance and typically set close to 1, say 0.95. Equation (4) is identical to

$$\frac{P_\theta(Y_{it_2} \geqslant \gamma, Y_{it_1} \geqslant \eta)}{P_\theta(Y_{it_1} \geqslant \eta)} = \frac{\int_\eta^\infty \int_\gamma^\infty \varphi_*(s_1, s_2) ds_2 ds_1}{\int_\eta^\infty \varphi_{Y_{it_1}}(s) ds} = q, \tag{5}$$

or

$$\frac{\int_\eta^\infty P_\theta(Y_{it_2} \geqslant \gamma | Y_{it_1} = s) \cdot \varphi_{Y_{it_1}}(s) ds}{P_\theta(Y_{it_1} \geqslant \eta)} = q, \tag{6}$$

where $\varphi_*(.)$ is the density function of the bivariate Gaussian (3), and $\varphi_{Y_{it_1}}$ is the density function of Y_{it_1}, i.e. $dN(a + bt_1, \sigma_\alpha^2 + \sigma_\beta^2 t_1^2 + \sigma_\varepsilon^2)$. Equation (5) has no apparent analytical solution but can be solved with a numerical algorithm easily. As for equation (6), it is known that $Y_{it_2} | Y_{it_1} = s \sim N(\mu_*, \sigma_*^2)$, where

$$\mu_* = (a + bt_2) - \frac{\sigma_\alpha^2 + \sigma_\beta^2 t_1 t_2}{\sigma_\alpha^2 + \sigma_\beta^2 t_1^2 + \sigma_\varepsilon^2}(a + bt_1 - s), \tag{7}$$

$$\sigma_*^2 = \sigma_\alpha^2 + \sigma_\beta^2 t_2^2 + \sigma_\varepsilon^2 - \frac{(\sigma_\alpha^2 + \sigma_\beta^2 t_1 t_2)^2}{\sigma_\alpha^2 + \sigma_\beta^2 t_1^2 + \sigma_\varepsilon^2}, \tag{8}$$

CP Formulation: Practicality

Lemma 3.1. *Suppose that (y_1, y_2) is bivariate normal with positive correlation. Then $\eta \mapsto P(y_2 \geqslant \gamma | y_1 \geqslant \eta)$ is increasing for any $\gamma \in \mathbb{R}$.*

Lemma 3.2. *If f and g are non-decreasing functions $f, g : \mathbb{R} \mapsto \mathbb{R}$, then $Cov(f(u), g(u)) \geqslant 0$, for any random variable u.*

Lemma 3.3. *Suppose that (y_1, y_2) is bivariate normal with positive correlation. Then $\eta \mapsto P(y_2 \geqslant \gamma | y_1 = \eta)$ is increasing, for any $\gamma \in \mathbb{R}$. Consequently $P(y_2 \geqslant \gamma | y_1 \geqslant \eta) \geqslant P(y_2 \geqslant \gamma | y_1 = \eta)$ for any γ, η.*

Remark 1: Correlation between (Y_{it_1}, Y_{it_2}) is most crucial in this setup

Remark 2: Target function (4) is lower bounded by $\mathrm{P}_\theta(Y_{it_2} \geq \gamma)$

Remark 3: No practical meaning if $q \geq \mathrm{P}_\theta\left(Y_{it_2} \geq \gamma \middle| Y_{it_1} \geq \gamma\right)$

Remark 4: We can solve η for large q while $\mathrm{P}_\theta(Y_{it_1} \geq \eta)$ is slim

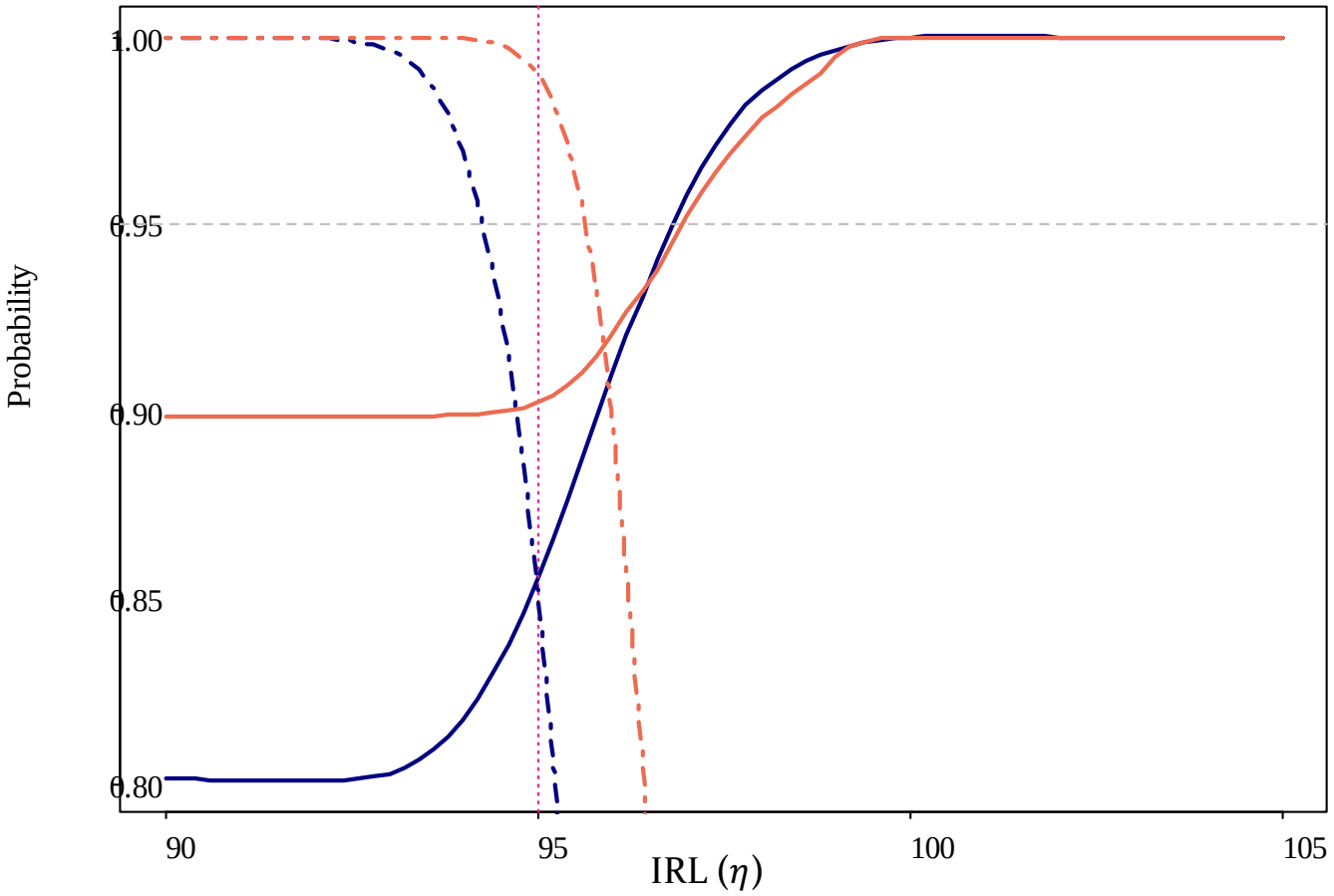
Remark 5: One may pursue the following instead, resulting IRL should be tighter

$$\eta = \arg\min_{\eta^*} P_\theta(Y_{it_2} \geqslant \gamma | Y_{it_1} = \eta^*) \geqslant q.$$

CP Formulation: Practicality

Red : a=97.2, b=-0.03, σ_β =0.03, σ_α=0.8, ρ=0.7, $p_0 = 99\%$, p_{24}=90%, $Cor(0,24) = 0.56$

Blue: a=96.2, b=-0.01, σ_β =0.03, σ_α=0.8, ρ=0.7, $p_0 = 90\%$, p_{24}=80%, $Cor(0,24) = 0.56$



NOVARTIS

CP Formulation: Estimation

Let Data be $x = \{x_{it}: i = 1,2, \dots; t = 0, t_1, \dots\}$ with batch index i and time index t

Frequentist: (Plug-in Estimator, consistency by mapping theorem)

(CP-Freq)

$$E\left[P_{\hat{\theta}(\mathbf{x})}(Y_{it_2} \geqslant \gamma | Y_{it_1} \geqslant \eta) \,|\mathbf{x}\right] \geqslant q, \qquad \text{for } \forall\theta.$$

wrt sample distribution of x

Bayesian:

$$E\left[P_{\theta}(Y_{it_2} \geqslant \gamma | Y_{it_1} \geqslant \eta) \,|\mathbf{x}\right] = q,$$

wrt posterior $p(\theta|x)$

The posterior is generally approximated by B samples $\{\theta_k\}_{k=1}^{B}$ via MCMC algorithm. *(CP-Bayes1)*

$$\frac{1}{B}\sum_{k=1}^{B} P_{\theta_k}(Y_{it_2} \geqslant \gamma | Y_{it_1} \geqslant \eta) = q.$$

Since B is a finite number, the above equation might not hold exactly for any η, but we can find the optimal η to as min $\left\{\eta^*: \frac{1}{B}\sum_{k=1}^{B} P_{\theta_k}(Y_{it_2} \geqslant \gamma | Y_{it_1} \geqslant \eta^*) \geqslant q\right\}$.

Alternatively, treat $\eta: \theta \mapsto \mathbb{R}^+$, get$\{\eta_k: \eta(\theta_k)\}_{k=1}^{B}$, and use posterior median or mean. We found posterior median is more robust to numerical issues. *(CP-Bayes2)*

Alternative: UCP Formulation

Assume analytical variance do not change. The stability trend is the only factor introduce difference between passing rate (of individuals w.r.t. LSS) at release and at EOSL

Alternative Objective:

Seek a IRL has at least the stringency as LSS at EOSL.

$$P_{\alpha_i,\beta_i}\left[\frac{P_\theta(Y_{i0} \geqslant \eta)}{P_\theta(Y_{iT} \geqslant \gamma)} \leqslant 1 \;\middle|\; \alpha_i, \beta_i\right] = q. \qquad (9)$$

It can be show this hints

$$\eta = \gamma - [b + \Phi^{-1}(1-q)\sigma_\beta]T,$$

where $\Phi^{-1}(1-q)$ denotes the $(1-q)^{th}$ quantile of the standard normal.

UCP Formulation: Estimation

Frequentist: (Plug-in Estimator, consistency by mapping theorem)

(UCP-Freq)

Bayesian: $\eta: \theta \mapsto \mathbb{R}^+$ given LSS (i.e. γ), b, σ_β, shelf life T and any desired probabilistic control q (>0.5). get $\{\eta_k: \eta(\theta_k)\}_{k=1}^{B}$, and use posterior median or mean. We found posterior median most robust due to numerical issues. *(UCP-Baye)*

We can also add an extra layer of probabilistic control, and use a posterior quantile as the point estimate.

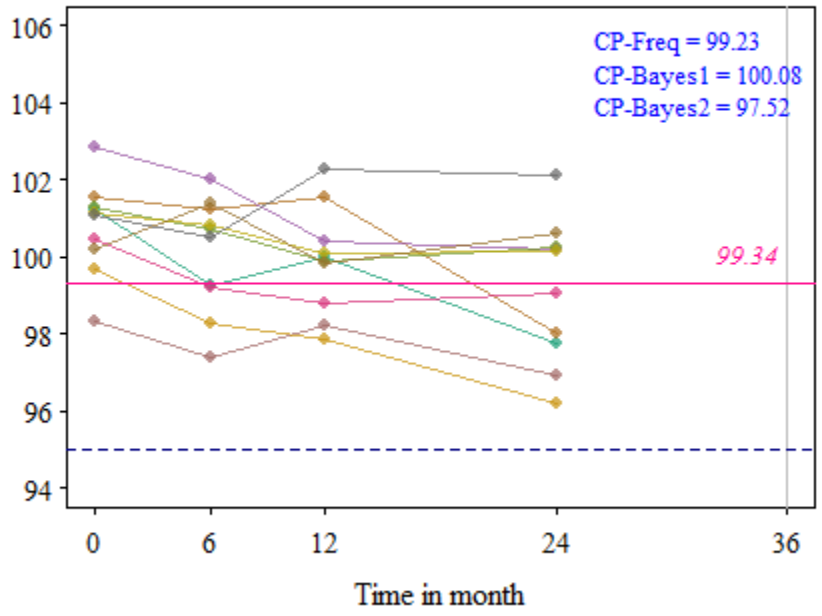
$$P\left(P_{\alpha_i,\beta_i}\left[\frac{P_\theta(Y_{i0} \geqslant \eta)}{P_\theta(Y_{iT} \geqslant \gamma)} \leqslant 1 \,\middle|\, \alpha_i, \beta_i\right] = q \,\middle|\, \mathrm{x}\right) = \xi.$$

This is simply a posterior quantile of η. When η is a lower-sided limit, we can take $\xi > 0.5$ to be conservative.

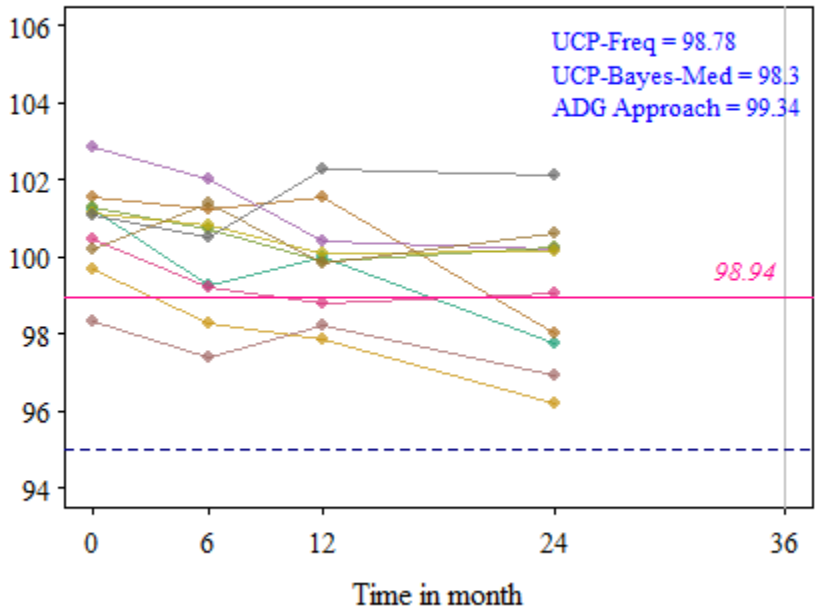
Simulated Example (1)

1. $P(Y_{i0}) = 0.9999, P(Y_{iT}) = 0.9, \sigma_\alpha = 1.2, \sigma_\beta = 0.03, \rho = 0.7, \sigma_\varepsilon^2 = (1-\rho)\sigma_\alpha^2/\rho$: This results in $a = 100.3$, $b = -0.084$, and a modest $\rho(0,T) = 0.56$.

CP Formulation



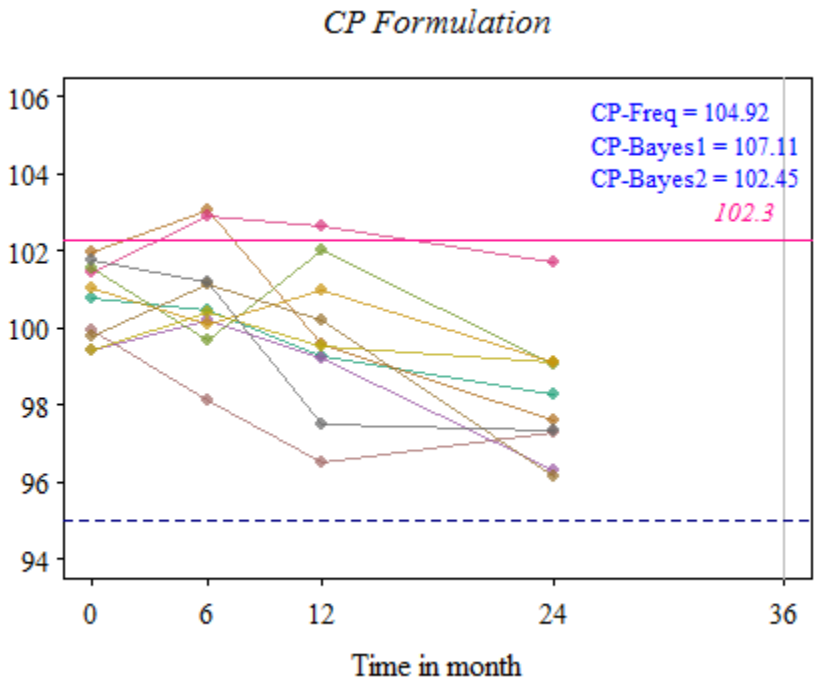
UCP Formulation

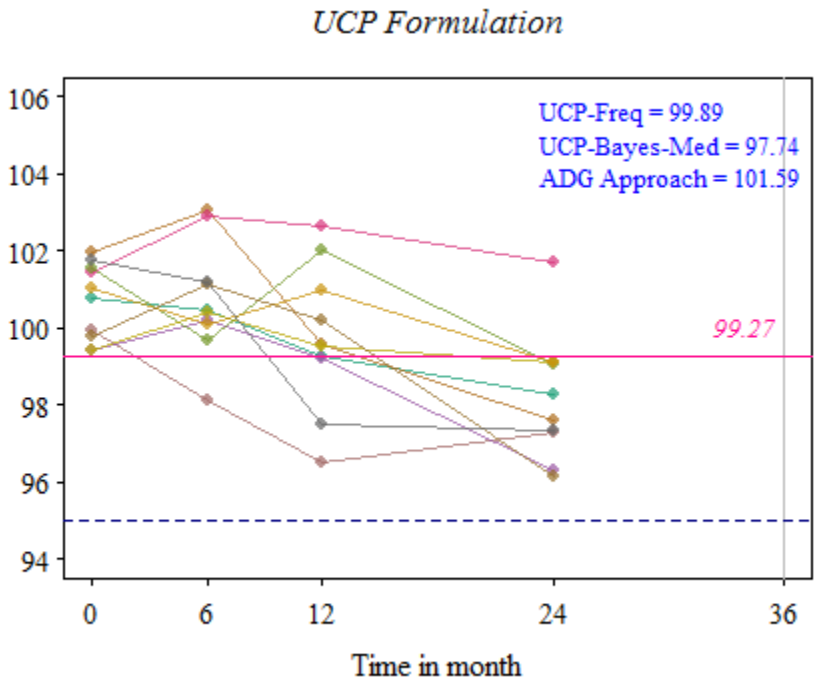


NOVARTIS

Simulated Example (2)

2. $P(Y_{i0}) = 0.9999, P(Y_{iT}) = 0.9, \sigma_\alpha = 1, \sigma_\beta = 0.06, \rho = 0.4, \sigma_\varepsilon^2 = (1-\rho)\sigma_\alpha^2/\rho$: This results in $a = 100.9$, $b = -0.068$, and a weak $\rho(0,T) = 0.24$.





THANK YOU