Measuring the dependence between endpoints for predictions across studies

V. Barboux
Supervisors: M. Pannaux & G. Saint-Hilary

ENSAI - Servier

November 06th 2017

Outline

Context and problematic

Context

- Principal endpoint of interest (**T**) is often difficult to assess : long follow-up, obtained with invasive painful techniques,...
- Surrogate endpoints (**S**): replacement endpoints able to detect if a tested drug gives convincing results easier/earlier

Problematic *How can we estimate the relationship between endpoints and its uncertainty, and then predict the treatment (**Z**) effect across studies using the different endpoints ?*

Note: We do not focus on approved validated surrogate endpoints

Surrogates' historic and definition

Fleming and DeMets (1996) : *"A correlate does not a surrogate make."*

Prentice (1989) :

- Z has a significant effect on S;
- Z has a significant effect on T;
- S has a significant effect on T;
- The treatment effect on T is fully captured by S, so (T|S,Z)=(T|S)

Z : treatment

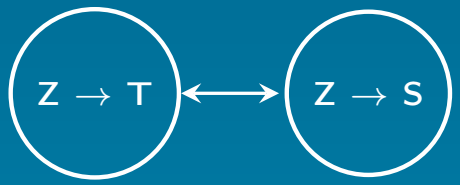
S : surrogate endpoint

T : true endpoint

$\Rightarrow$ The fourth criterion is often unverifiable (except for 2 binary endpoints) and proposed solutions (Proportion Explained, Relative Effect,...) are not convincing.

Meta-analytic approach

Buyse & Molenbergh (1998)

- 

- Main approach to get a prediction model
- Take into account a large collection of trial
- Any range of treatments' classes can be considered
- Multicenter trial & Meta-analysis
- Validity of a surrogate $\approx$ quality of the prediction

Two models

- Individual patient data (IPD)
- Aggregated data (AD)

Model: Individual Level

We consider normal endpoints.

For a patient j $\in [\![1, n_i]\!]$ and a trial i $\in [\![1, N]\!]$:

$$\begin{pmatrix} S_{ij} \\ T_{ij} \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} \xi_{S_i} + \mu_{S_i} Z_{ij} \\ \xi_{T_i} + \mu_{T_i} Z_{ij} \end{pmatrix}, \Sigma \right].$$

With ξ_{S_i} and ξ_{T_i} the intercepts and μ_{S_i} and μ_{T_i} the true treatment effects of Z respectively on S and T for the trial i. The variance-covariance matrix Σ is given by :

$$\Sigma = \begin{pmatrix} \tau_S^2 & \rho_{ST}\tau_S\tau_T \\ & \tau_T^2 \end{pmatrix}.$$

Where ρ_{ST} is the within-trial correlation between S_{ij} and T_{ij}. τ_S and τ_T are the within-trial standard deviation for S_{ij} and T_{ij} respectively.

Model: Trial Level

For each trial i:

$$\begin{pmatrix} \xi_{S_i} \\ \xi_{T_i} \\ \mu_{S_i} \\ \mu_{T_i} \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} \alpha_S \\ \alpha_T \\ \beta_S \\ \beta_T \end{pmatrix}, D \right].$$

With:

$$D = \begin{pmatrix} \tau^2_{\alpha_S} & \rho_{\alpha_S\alpha_T}\tau_{\alpha_S}\tau_{\alpha_T} & \rho_{\alpha_S\beta_S}\tau_{\alpha_S}\tau_{\beta_S} & \rho_{\alpha_S\beta_T}\tau_{\alpha_S}\tau_{\beta_T} \\ & \tau^2_{\alpha_T} & \rho_{\alpha_T\beta_S}\tau_{\alpha_T}\tau_{\beta_S} & \rho_{\alpha_T\beta_T}\tau_{\alpha_T}\tau_{\beta_T} \\ & & \tau^2_{\beta_S} & \rho_{\beta_S\beta_T}\tau_{\beta_S}\tau_{\beta_T} \\ & & & \tau^2_{\beta_T} \end{pmatrix}$$

With :

- $\tau_.$ the between-trial standard deviation
- $\rho_{..}$ the between-trial correlation between parameters

Prediction

Suppose that we have fitted the previous mixed-effect model on learning data from $i \in [\![1, n]\!]$ trial, and we have data on the surrogate from one new trial i=0.

For $i = 0$ and $j \in [\![1, n_0]\!]$, we can fit the model for the surrogate endpoint:

$$S_{0j} = \xi_{S0} + \mu_{S_0} Z_{0j} + \epsilon_{S0j}$$

We can predict μ_{T_0} given $\hat{\xi}_{S_0}$, $\hat{\mu}_{S_0}$ and $\hat{\vartheta} = (\ \hat{\beta}_T,\ \hat{\alpha}_S,\ \hat{D}[1,1],\ \hat{D}[1,3],\ \hat{D}[1,4],\ \hat{D}[3,3],\ \hat{D}[3,4])$.

Quality of the prediction

Simplified writing of the trial-specific treatment effect prediction model:

$$\mu_{T_0} = A + B\mu_{S_0} + \epsilon$$

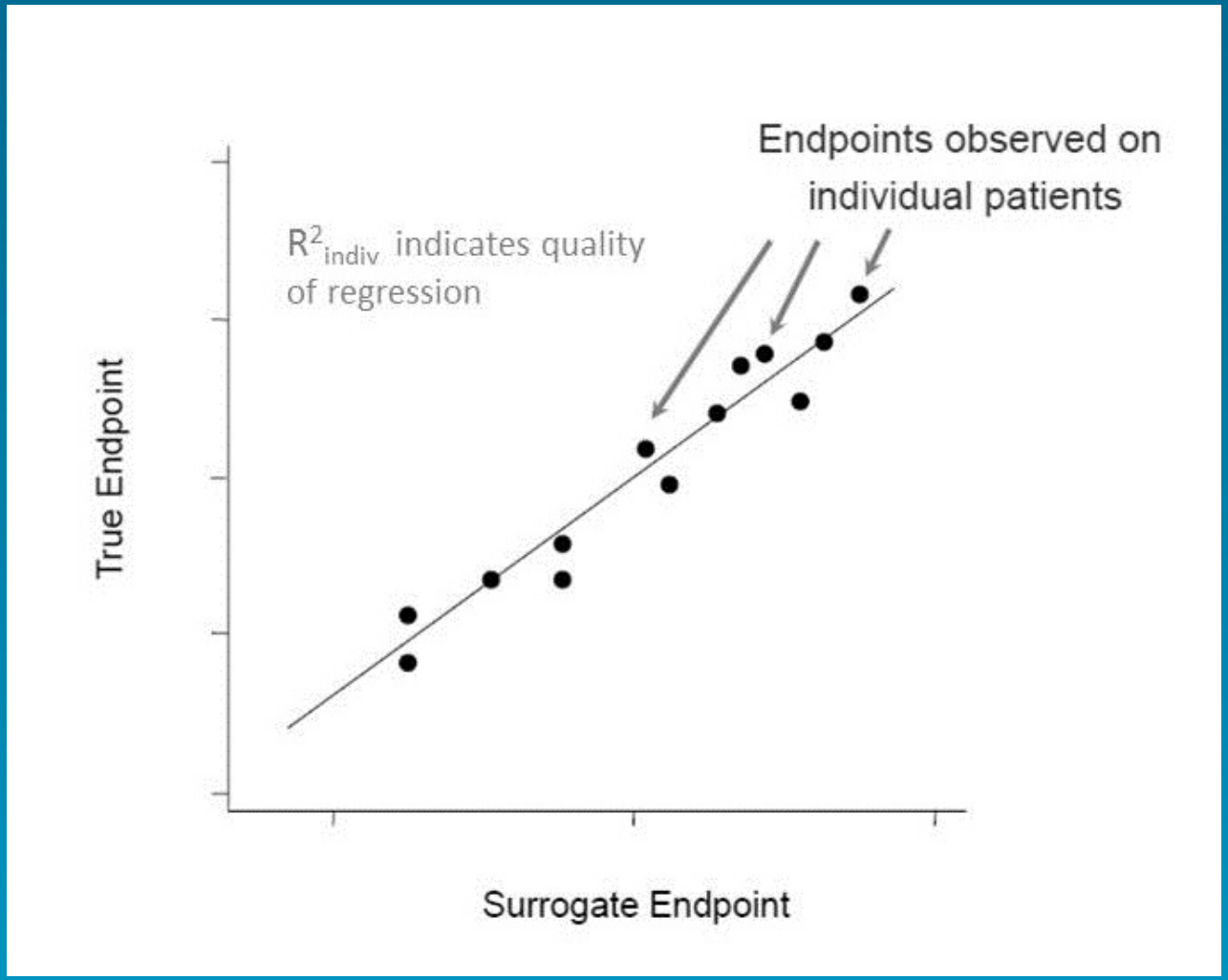
$$Var(\hat{\mu_{T_0}} - \mu_{T_0}) \approx \underbrace{f\left[\mathbb{V}(\hat{\mu_{S_0}})\right]}_{(1)} + \underbrace{f\left[\mathbb{V}(\hat{A}, \hat{B})\right]}_{(2)} + \underbrace{(1 - \mathcal{R}^2_{trial})\tau^2_{\beta_T}}_{(3)}.$$

(1) The sampling variance of the new trial on the surrogate endpoint. (Small if large size of the new trial).

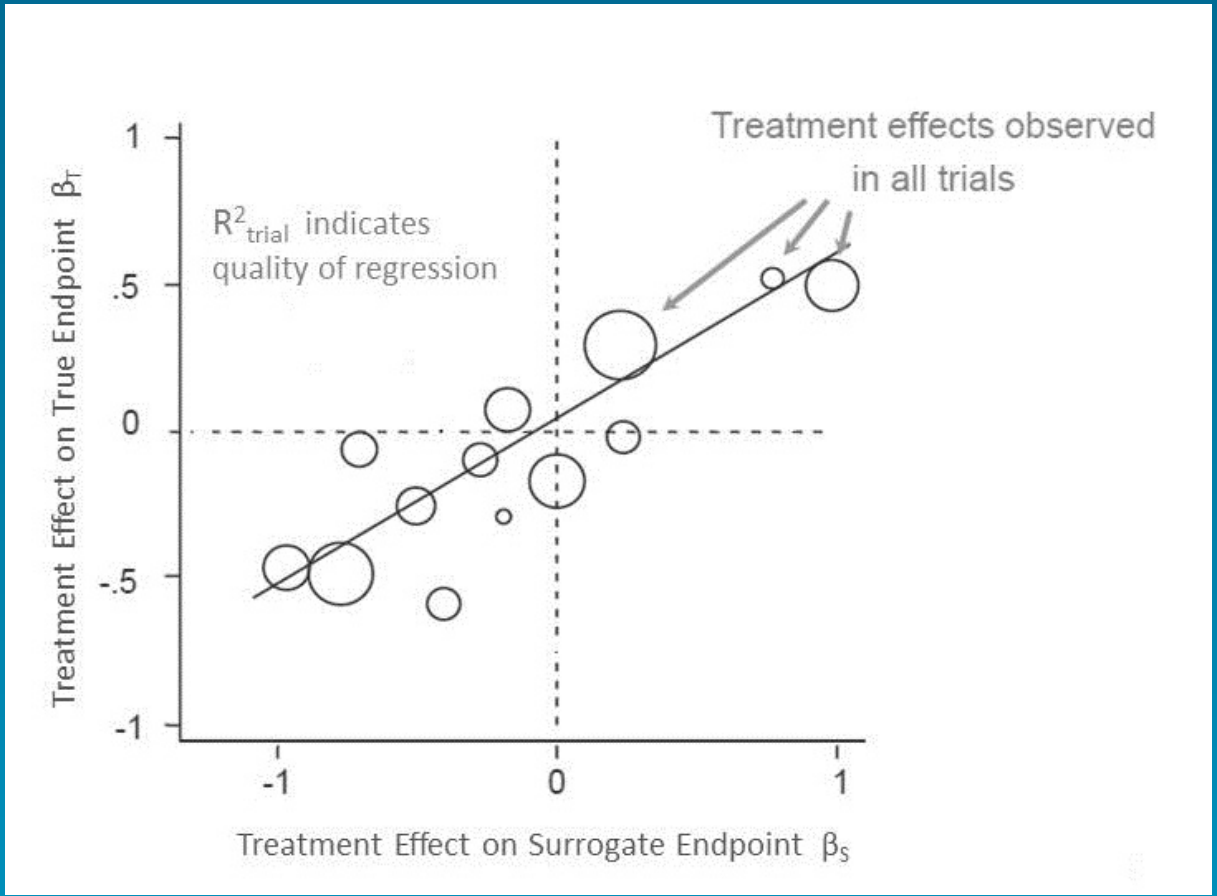
(2) The sampling variance of the past data. (Small if meta-analysis is large).

(3) The dependence between the treatment effects on the surrogate and the true endpoints. (Small if: the surrogate is valid i.e. good predictor).

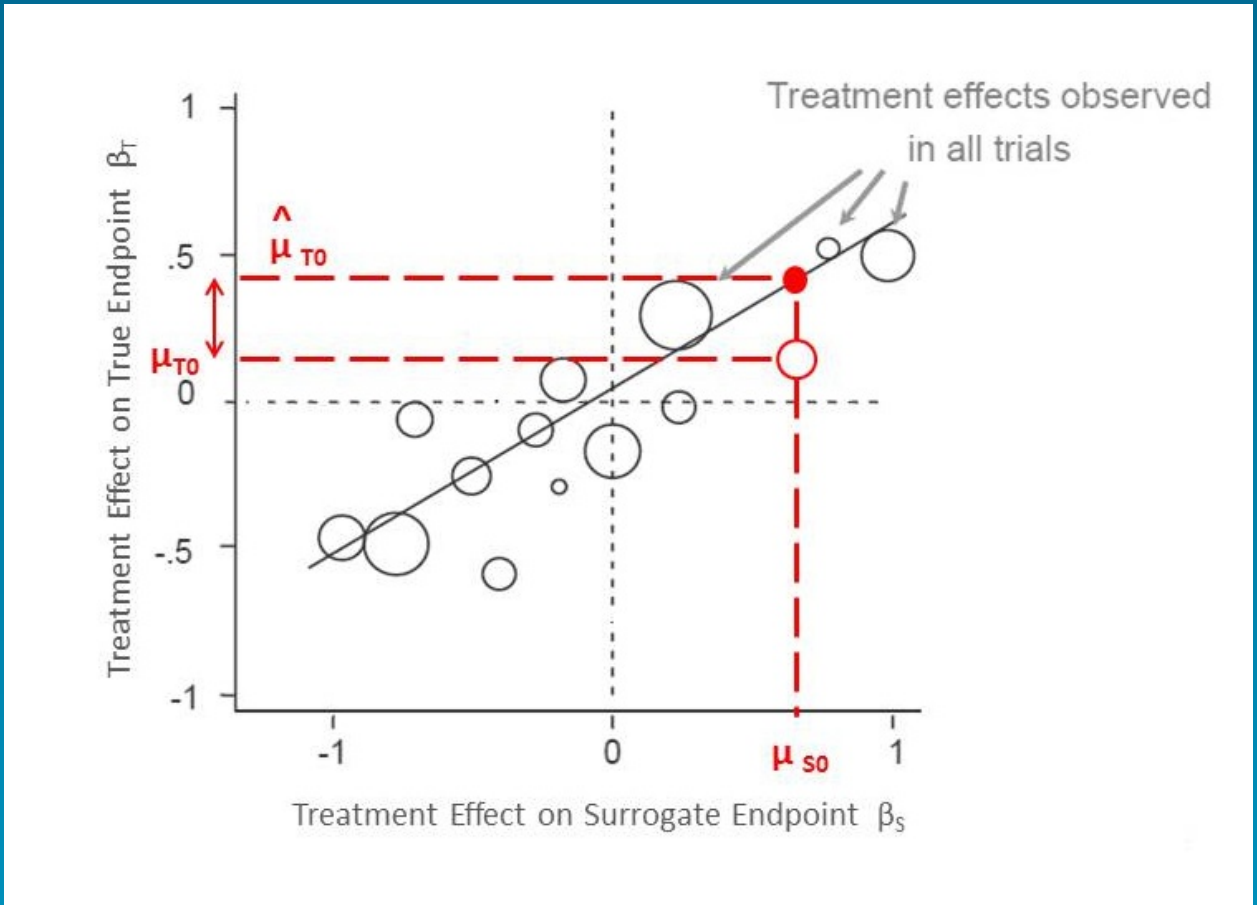
Prediction of T from S



Prediction of treatment effect



Prediction of treatment effect



Some quantification

Relative Bias (%) : $(\frac{\bar{\mathcal{R}}^2}{\mathcal{R}^2} - 1)100$ with $\bar{\mathcal{R}}^2 = \sum_i \bar{\mathcal{R}}_i^2$, $\bar{\mathcal{R}}_i^2$ representing the posterior mean in the ith simulation of $\mathcal{R}^2$.

Surrogate Threshold Effect (STE) : the smallest treatment effect for which S predict a significant treatment effect on T.

Simulation

We defined:

- $\alpha_S = 10$ and $\alpha_T = 20$;
- $\beta_S = 10$ and $\beta_T = 5$;
- $D = \sigma^2 \begin{pmatrix} 1 & 0.8 & 0 & 0 \\ & 1 & 0 & 0 \\ & & 1 & \rho \\ & & & 1 \end{pmatrix}$, with $\sigma^2 = 15$;
- We conduct the analyses for $\rho^2 = \mathcal{R}^2_{trial} = (0.7, 0.8, 0.9)$;
- $\Sigma = \begin{pmatrix} 1 & \rho_2 \\ & 1 \end{pmatrix}$ with $\rho_2^2 = \mathcal{R}^2_{indiv} = 0.7$.

Simulation

$\mathcal{R}^2_{indiv}$ \ $\mathcal{R}^2_{trial}$	0.7	0.8	0.9	Estimated values
0.7	0.889	0.932	0.967	ρ
	0.707	0.707	0.707	$\hat{\mathcal{R}}^2_{indiv}$
	0.791	0.869	0.936	$\hat{\mathcal{R}}^2_{trial}$
	3.864	4.657	4.852	STE
	12.230 %	7.160 %	3.319 %	Relative bias (trial level)
	<0.001 %	<0.001 %	<0.001 %	Relative bias (individual level)

Table 1: Simulation results with three different values of $\hat{\mathcal{R}}^2_{trial}$

Case presentation

Sormani et al. (2009)

- Multiple Sclerosis (MS)
- Number of brain lesions using Magnetic Resonance Imagning (MRI), as surrogate
- Relapse Rate (REL), as true endpoint

Sormani et al. (2013)

- Relapse Rate (REL), as surrogate
- Disability progression (DIS), as true endpoint

Updated data (2013-2017 trials)

$\Rightarrow$ Predictive model

Considered treatment effect

MRI the risk ratio (RR) based on MRI lesion counts;

REL the risk ratio (RR) associated with annualized relapse rate;

DIS the disability progression is assessed by the Expanded Disability Status Scale (EDSS) score, both risk ratio (RR) based on the proportion of patients with disability progression and hazard ratio (HR) based on the time to progression are considered.

$\Rightarrow$ We considered multiple surrogates

Available trials

	Short term (6-9 months)	Long term (12-36 months)	Total
2 arms	14	12	26
3 arms	15	8	23
4 arms	2	4	6
5 arms	0	1	1
Total	31	25	56

Table 2: Repartition of the trials per arm and phase.

$\Rightarrow$ We considered multiple trial-arms

Within-trial level

The joint distribution, for trial i, for the within-trial level is given by:

$$\begin{pmatrix} \hat{\mu}_{T_i} \\ \hat{\mu}_{S_i} \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} \mu_{T_i} \\ \mu_{S_i} \end{pmatrix}, \Omega_i \right].$$

With:

$$\Omega_i = \begin{pmatrix} \sigma^2_{Ti} & \rho_{STi}\sigma_{Si}\sigma_{Ti} \\ & \sigma^2_{Si} \end{pmatrix}$$

- ρ_{STi} the correlation in trial i between the estimated treatment differences conditional on μ_{T_i} and μ_{S_i}, and it is called within-trial correlation. (Unknown);
- μ_{S_i} and μ_{T_i} the means true treatment effects respectively on the surrogate and the true endpoint in trial i;

Between-trial level

We have the between trial level:

$$\begin{pmatrix} \mu_{Ti} \\ \mu_{Si} \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} \beta_T \\ \beta_S \end{pmatrix}, \Lambda \right].$$

With :

$$\Lambda = \begin{pmatrix} \tau^2_{\beta_T} & \rho_{\beta_T \beta_S} \tau_{\beta_T} \tau_{\beta_S} \\ & \tau^2_{\beta_S} \end{pmatrix}$$

- β_T and β_S the mean true treatment effects on the true endpoint and the surrogate endpoint;
- $\tau^2_{\beta_T}$ and $\tau^2_{\beta_S}$ the between trial variance corresponding to each true effect μ_{Ti} and μ_{Si};
- $\rho_{\beta_T \beta_S}$ the between trial correlation between μ_{Ti} and μ_{Si}.

Within-trial level (rewriting)

$\hat{\mu}_{T_i}|\hat{\mu}_{S_i}$ follows a normal distribution with the univariate conditional distributions:

$$\begin{aligned}\mu_{S_i} &\sim \mathcal{N}(\eta_S, \psi_S^2)\\ \eta_{T_i} &= \lambda_1 + \lambda_2 \mu_{S_i}\\ \mu_{T_i}|\mu_{S_i} &\sim \mathcal{N}(\eta_{T_i}, \psi_T^2)\end{aligned}$$

And then we can re-write the model as:

$$\begin{pmatrix}\hat{\mu}_{Ti}\\ \hat{\mu}_{Si}\end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix}\lambda_1 + \lambda_2\mu_{Si}\\ \mu_{S_i}\end{pmatrix}, \begin{pmatrix}\tau^2_{\beta_T} - \lambda_2\tau^2_{\beta_S} + \sigma^2_{Ti} & \rho_{STi}\sigma_{Si}\sigma_{Ti}\\ & \sigma^2_{Si}\end{pmatrix}\right].$$

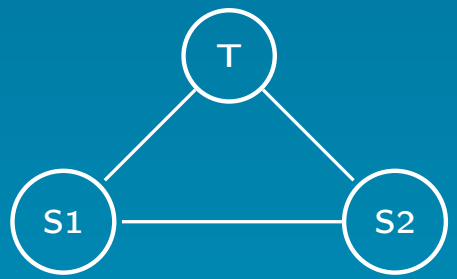
Multiple surrogates

Extended within-trial model:

$$\begin{pmatrix} \hat{\mu}_{1i} \\ \hat{\mu}_{2i} \\ \hat{\mu}_{3i} \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} \mu_{1i} \\ \mu_{2i} \\ \mu_{3i} \end{pmatrix}, \Omega_i = \begin{pmatrix} \sigma^2_{1i} & \rho_{12i}\sigma_{1i}\sigma_{2i} & \rho_{13i}\sigma_{1i}\sigma_{2i} \\ & \sigma^2_{2i} & \rho_{23i}\sigma_{2i}\sigma_{3i} \\ & & \sigma^2_{3i} \end{pmatrix} \right].$$

Multiple scenario:

(1)



(2)

Multiple arms

With only one surrogate endpoint and one final endpoint, the within-trial level can be defined with a multivariate normal model, for trial i:

$$\begin{pmatrix} \hat{\mu}_{T_1 i} \\ \hat{\mu}_{T_2 i} \\ \hat{\mu}_{S_1 i} \\ \hat{\mu}_{S_2 i} \end{pmatrix} \sim \mathcal{N} \left[\begin{pmatrix} \mu_{T_1 i} \\ \mu_{T_2 i} \\ \mu_{S_1 i} \\ \mu_{S_2 i} \end{pmatrix}, \begin{pmatrix} \sigma^2_{T_1 i} & \rho_{Ti}\sigma_{T_1 i}\sigma_{T_2 i} & \rho_{11i}\sigma_{T_1 i}\sigma_{S_1 i} & \rho_{12i}\sigma_{T_1 i}\sigma_{S_2 i} \\ & \sigma^2_{T_2 i} & \rho_{21i}\sigma_{T_2 i}\sigma_{S_1 i} & \rho_{22i}\sigma_{T_2 i}\sigma_{S_2 i} \\ & & \sigma^2_{S_1 i} & \rho_{Si}\sigma_{S_1 i}\sigma_{S_2 i} \\ & & & \sigma^2_{S_2 i} \end{pmatrix} \right]$$

With ρ_{Ti} and ρ_{Si} within-arm correlations.

Application

List of tested model:

Model	Follow-up	Surrogate(s) endpoint	Final endpoint
1	All	MRI	Relapse
2	Short term (6-9m)	MRI	Relapse
3	Long term (12-36m)	MRI	Relapse
4	Long term (12-36m)	Relapse	Disability (RR)
5	Long term (12-36m)	Relapse	Disability (HR)
6	Long term (12-36m)	MRI + Relapse	Disability (RR)
7	Long term (12-36m)	MRI + Relapse	Disability (HR)

Table 3: List of tested AD models in RRMS

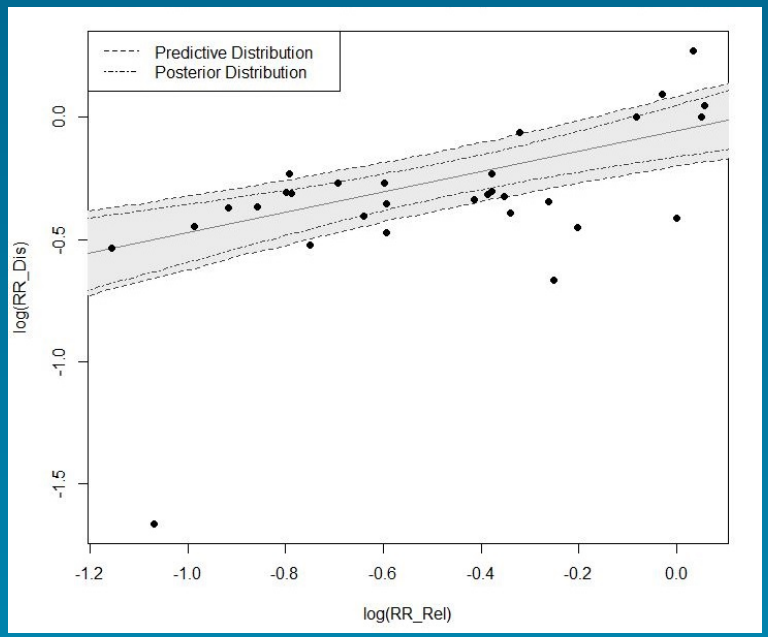
In the litterature : $\rho = 0.05$.
We will consider model 4 and 6 in the rest of the section.

Results

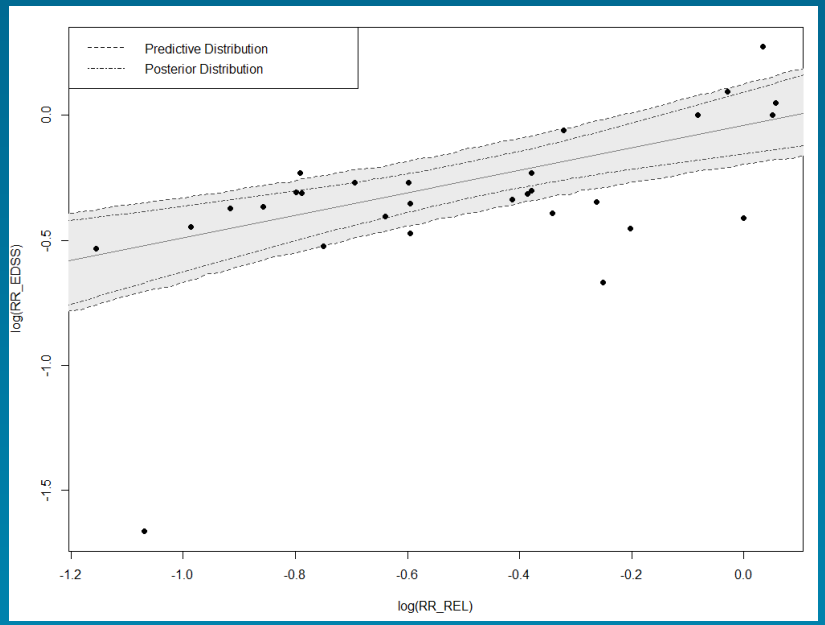
Model	Parameters	ρ=0.05 [95%CrI]
Two level (4)	α_1	-0.056 [-0.159 ;0.053]
	β_1	0.415 [0.237;0.578]
	τ	0.039 [0.005 ; 0.090]
Three level (6)	α_1	-0.038 [-0.163;0.083]
	β_1	0.453 [0.251 ;0.661]
	α_2	0.028 [-0.128; 0.175]
	β_2	0.508 [0.379 ;0.638]
	τ	0.044 [0.004 ;0.098]
	ω	0.101 [0.025;0.197]

Table 4: Posterior means and 95% credible intervals for two-level and three-level models. We consider the posterior distributions for the following regressions : log(DIS RR)$= \alpha_1 + \beta_1$log(REL)$+\epsilon$ with $\epsilon \sim \mathcal{N}(0, \tau^2)$ and for the three-level model only log(REL)$= \alpha_2 + \beta_2$log(MRI)$+\zeta$ with $\zeta \sim \mathcal{N}(0, \omega^2)$

Results



(a) $\rho = 0.05$, two-level



(b) $\rho = 0.05$, three-level

Figure 1: Fitted two-level and three-level model for log(DIS RR) based on log(REL). The three-level model takes into account log(MRI). Are represented: the posterior distribution of the regression line of log(DIS RR) on log(REL) with a 95% credible interval and the 95% predictive interval.

Discussion

Bayesian approach :

- choice of prior => subjectivity
- no clear guidance for the choice of prior

IPD model :

- find adequate model for dependence between endpoints at the individual level
- availability of individual patient data

AD model :

- within-study correlation usually not available
- several scenarios to consider about association between treatment effects

Meta-analytic approach: data demanding

Conclusion

There are other methods to assess the validity of surrogate endpoints. Meta-analytic approaches have the advantage assess the quality of the prediction & to provide prediction model.

Other approach: Causal Inference

- does not provide practical prediction model
- lot of investigation are made on this (should be follow-up)
- strong assumptions needed

Thank you for your attention.

Any questions ?