Web Appendix to 'Univariate versus multivariate surrogate endpoints in the single-trial setting.'

Part I

Proof of Lemma's 1 and 2.

Proof of Lemma 1. In the following Lemma 1 will be proven.

1. R_H^2 *is invariant by bijective transformations of* ΔT *and* $\boldsymbol{\Delta S}$
The result follows from the fact that the mutual information $I\left(\Delta T, \boldsymbol{\Delta S}\right)$ is invariant by bijective transformations of ΔT and $\boldsymbol{\Delta S}$.

2. $0 \leq R_H^2 \leq 1$
It is a direct consequence of $R_H^2 = \max_{\mathbf{t}} \left[\text{corr}\left(\Delta T, \boldsymbol{t}' \boldsymbol{\Delta S}\right)\right]^2$.

3. $R_H^2 = 0$ *if and only if* $\sigma_{\Delta S_k \Delta T} = 0$ *for all* $r = 1, 2 \ldots p$
Given that $\boldsymbol{\Sigma}_{\Delta S}^{-1}$ is positive-definite it follows from expression (4) in the manuscript that $R_H^2 = 0$ if and only if $\boldsymbol{\Sigma}_{\boldsymbol{\Delta S}\Delta T} = \mathbf{0}$. Moreover $\boldsymbol{\Sigma}_{\boldsymbol{\Delta S}\Delta T} = \mathbf{0}$ if and only if $\sigma_{\Delta S_r \Delta T} = 0$ for all $r = 1, 2 \ldots p$.

4. $R_H^2 = 1$ *if and only if there exists a deterministic relationship between* ΔT *and* $\boldsymbol{\Delta S}$.
From $R_H^2 = \max_{\mathbf{t}} \left[\text{corr}\left(\Delta T, \boldsymbol{t}' \boldsymbol{\Delta S}\right)\right]^2$ it is clear that $R_H^2 = 1$ if and only if there exists a $\boldsymbol{t}_*$.

□

Proof of Lemma 2. In the following Lemma 2 will be proven.

Notice that for every vector $\boldsymbol{t} \in \mathbb{R}^p$ there exists a $\boldsymbol{t}_0 = (\boldsymbol{t}, 0)' \in \mathbb{R}^{p+1}$ so that $\text{corr}\left(\Delta T, \boldsymbol{t}'\boldsymbol{\Delta S}\right) = \text{corr}\left(\Delta T, \boldsymbol{t}_0'\boldsymbol{\Delta S}_*\right)$. Consequently,

$$R_H^2 = \max_{\boldsymbol{t}\in\mathbb{R}^p}\left[\text{corr}\left(\Delta T, \boldsymbol{t}'\boldsymbol{\Delta S}\right)\right]^2 = \max_{\boldsymbol{t}_0\in\mathbb{R}^{p+1}}\left[\text{corr}\left(\Delta T, \boldsymbol{t}_0'\boldsymbol{\Delta S}_*\right)\right]^2 \leq \max_{\boldsymbol{t}_*\in\mathbb{R}^{p+1}}\left[\text{corr}\left(\Delta T, \boldsymbol{t}_*'\boldsymbol{\Delta S}_*\right)\right]^2 = R_{H*}^2$$

□

Part II

Case study analysis using the R package *Surrogate*

1 The dataset: the transPAT microbiome intervention study

Data from the transPAT experiment (Ruiz *et al.*, 2017) are used to illustrate the multivariate surrogate evaluation methodology. TransPAT is an animal study that was conducted to evaluate the influence of an antibiotic treatment on the immune system (Immunoglobulin A level, IgA level) and the microbiome of an animal. The microbiome is composed of a wide variety of operational taxonomic units (OTUs), which are essentially proxies for microbial species. The transPAT dataset contains information of $N = 15$ germ-free mice that received cecal contents of a donor mouse. The cecal contents of the donor mouse was either exposed (experimental treatment) or not exposed (control treatment) to a tylosin pulse. A total of $n = 7$ and $n = 8$ mice received the experimental and the control treatments, respectively. The relative abundance of a total of 67 OTUs was assessed at day 12 of the experiment. In the current analysis, it will be evaluated whether the individual causal treatment effect on the relative abundance of one or more of the 67 OTUs at day 12 of the experiment (ΔS_1, ΔS_2, ..., ΔS_{67}) is predictive for the individual causal treatment effect on IgA level at day 20 of the experiment (ΔT). The data are available on github (see `https://github.com/blaser-lab/Paper-Ruiz-2017`).

2 The multivariate individual causal association (R^2_H)

The function `ICA.ContCont.MultS()` in the *Surrogate* package allows for computing the individual causal association (ICA) in the setting where both the true and surrogate endpoints are Continuous normally distributed variables, and where Multiple Surrogates are considered. This function computes the multivariate individual causal association (R^2_H) following the procedure that was detailed in Van der Elst *et al.* (2018). The function requires the user to specify the following arguments:

- `M=` : the number of multivariate individual causal association (R^2_H) values that have to be sampled by the algorithm. Default `M=500`.
- `N=` : the number of patients/subjects in the dataset.
- `Sigma=` : the variance-covariance matrix $\mathbf{\Sigma}$ of the true and surrogates endpoints in both treatment

groups (i.e., T_0, T_1, S_{10}, S_{11}, ..., S_{p0}, S_{p1}).

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{T_0T_0} & \sigma_{T_0T_1} & \sigma_{T_0S_{10}} & \sigma_{T_0S_{11}} & \sigma_{T_0S_{20}} & \sigma_{T_0S_{21}} & \cdots & \sigma_{T_0S_{p0}} & \sigma_{T_0S_{p1}} \\ \sigma_{T_0T_1} & \sigma_{T_1T_1} & \sigma_{T_1S_{10}} & \sigma_{T_1S_{11}} & \sigma_{T_1S_{20}} & \sigma_{T_1S_{21}} & \cdots & \sigma_{T_1S_{p0}} & \sigma_{T_1S_{p1}} \\ \sigma_{T_0S_{10}} & \sigma_{T_1S_{11}} & \sigma_{S_{10}S_{10}} & \sigma_{S_{10}S_{11}} & \sigma_{S_{10}S_{20}} & \sigma_{S_{10}S_{21}} & \cdots & \sigma_{S_{10}S_{p0}} & \sigma_{S_{10}S_{p1}} \\ \sigma_{T_0S_{11}} & \sigma_{T_1S_{11}} & \sigma_{S_{10}S_{11}} & \sigma_{S_{11}S_{11}} & \sigma_{S_{11}S_{20}} & \sigma_{S_{11}S_{21}} & \cdots & \sigma_{S_{11}S_{p0}} & \sigma_{S_{11}S_{p1}} \\ \sigma_{T_0S_{20}} & \sigma_{T_1S_{20}} & \sigma_{S_{10}S_{20}} & \sigma_{S_{11}S_{20}} & \sigma_{S_{20}S_{20}} & \sigma_{S_{20}S_{21}} & \cdots & \sigma_{S_{20}S_{p0}} & \sigma_{S_{20}S_{p1}} \\ \sigma_{T_0S_{21}} & \sigma_{T_1S_{21}} & \sigma_{S_{20}S_{21}} & \sigma_{S_{11}S_{21}} & \sigma_{S_{20}S_{21}} & \sigma_{S_{21}S_{21}} & \cdots & \sigma_{S_{21}S_{p0}} & \sigma_{S_{21}S_{p1}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \sigma_{T_0S_{p0}} & \sigma_{T_1S_{p0}} & \sigma_{S_{10}S_{p0}} & \sigma_{S_{11}S_{p0}} & \sigma_{S_{20}S_{p0}} & \sigma_{S_{21}S_{p0}} & \cdots & \sigma_{S_{p0}S_{p0}} & \sigma_{S_{p0}S_{p1}} \\ \sigma_{T_0S_{p1}} & \sigma_{T_1S_{p1}} & \sigma_{S_{10}S_{p1}} & \sigma_{S_{11}S_{p1}} & \sigma_{S_{20}S_{p1}} & \sigma_{S_{21}S_{p1}} & \cdots & \sigma_{S_{p0}S_{p1}} & \sigma_{S_{p1}S_{p1}} \end{pmatrix}. \quad (1)$$

The unidentifiable covariances in $\boldsymbol{\Sigma}$ should be given the value `NA`.

- `G=` : a vector of values that should be considered for the unidentified correlations. Default `G=seq(-1, 1, by=0.00001)`.
- `Seed=` : the seed to be used in the analysis (for reproducibility).
- `Show.Progress=` : a logical indicator that can be used to request visual feedback regarding the progress of the algorithm in finding R^2_H values. When `Show.Progress=TRUE` is used, 1% done..., 2% done..., etc is shown in the R console when the function is running. This option is mainly useful when 4 or more surrogates are being considered in the analysis, because finding positive definite $\boldsymbol{\Sigma}$ (which are needed to compute R^2_H) can take a very long time when the dimensionality of $\boldsymbol{\Sigma}$ is high (see Van der Elst *et al.*, 2018).

Identification of subset of good surrogates A hierarchical approach with forward selection was used to identify the best set of surrogates. To this end, univariate analyses were conducted for each of the $p = 67$ candidate surrogates. The candidate univariate surrogate that had the highest median R^2_H was retained. In the second step, multivariate analyses were conducted that included the first identified candidate surrogate combined with a second one (i.e., one of the $p - 1 = 66$ remaining OTUs). Again, for the 66 analyses conducted, the bivariate surrogate that led to the highest median R^2_H was retained. This procedure was repeated until $\boldsymbol{\Delta S}$ had a median $R^2_H \geq 0.90$. The so-obtained final vector of surrogates included S_1 = OTU 44, S_2 = OTU 1 and S_3 = OTU 59, with variance-covariance matrix:

$$\boldsymbol{\Sigma} = \begin{pmatrix} 176.7790691375 & NA & 0.0936263266 & NA & -1.5311587825 & NA & 0.0000404661 & NA \\ NA & 11.8330574635 & NA & -0.0002432784 & NA & -0.4853893940 & NA & -0.0018366108 \\ 0.0936263266 & NA & 0.0000621257 & NA & -0.0006507995 & NA & -0.0000002111 & NA \\ NA & -0.0002432784 & NA & 0.0000002061 & NA & 0.0000572090 & NA & 0.0000002735 \\ -1.5311587825 & NA & -0.0006507995 & NA & 0.0238066268 & NA & -0.0000025912 & NA \\ NA & -0.4853893940 & NA & 0.0000572090 & NA & 0.0487158192 & NA & 0.0002031266 \\ 0.0000404661 & NA & -0.0000002111 & NA & -0.0000025912 & NA & 0.0000000065 & NA \\ NA & -0.0018366108 & NA & 0.0000002735 & NA & 0.0002031266 & NA & 0.0000008971 \end{pmatrix}.$$

The following code can be used to request a surrogacy analysis for the transPAT case study, using all three surrogates and the default options of the `ICA.ContCont.MultS()` function (runtime about 2 hours):

```
# First define the Sigma matrix for T and S_1 = OTU 44, S_2 = OTU 1 and S_3 = OTU 59
> Sigma <- matrix(data=c(
1.767791e+02, NA, 9.362633e-02, NA, -1.531159e+00, NA, 4.046615e-05, NA,
NA, 11.8330574635, NA, -2.432784e-04, NA, -4.853894e-01, NA, -1.836611e-03,
9.362633e-02, NA, 6.212567e-05, NA, -6.507995e-04, NA, -2.111307e-07, NA,
NA, -0.0002432784, NA, 2.060978e-07, NA, 5.720905e-05, NA, 2.734616e-07,
-1.531159e+00, NA, -6.507995e-04, NA, 2.380663e-02, NA, -2.591242e-06, NA,
NA, -0.4853893940, NA, 5.720905e-05, NA, 4.871582e-02, NA, 2.031266e-04,
4.046615e-05, NA, -2.111307e-07, NA, -2.591242e-06, NA, 6.545024e-09, NA,
NA, -0.0018366108, NA, 2.734616e-07, NA, 2.031266e-04, NA, 8.970605e-07
), nrow=8)

# Conduct the analysis
> ICA_S1S2S3 <- ICA.ContCont.MultS(Sigma = Sigma, N=15)
```

The fitted object `ICA_S1S2S3` of class `ICA.ContCont.MultS` contains the results. The `names()` function can be applied to the fitted object to obtain a list of the components in this object:

```
> names(ICA_S1S2S3)

# Generated output:

"R2_H"                  "Corr.R2_H"             "Lower.Dig.Corrs.Sigma" "Call"
```

These components contain the following information:

- `R2_H=` : the R^2_H values, computed as detailed in Van der Elst *et al.* (2018).
- `Corr.R2_H=` : the R^2_H values correspond to the coefficient of determination of a multiple regression model in which ΔT is regressed on ΔS_1, ΔS_2, ..., and ΔS_k. As a consequence, R^2_H can never decrease when additional surrogates are added to the existing ones. One could argue that this property of R^2_H is less desirable, because a more parsimonious model (i.e., a model that uses less surrogates) may be preferred over a less parsimonious model (i.e., a model that uses more surrogates) if both models explain about the same amount of variance in ΔT. In addition, the sample coefficient of determination of a regression model is a positively biased estimate of the population coefficient of determination (Cohen *et al.*, 2003). This bias is small when only a small number of surrogates is used and when the sample size is large, but it becomes more substantial when these conditions are not fulfilled. Several proposals have been made to correct for this bias (Yin and Fan, 2001). The often-used Wherry's equation makes the correction $1 - \left(1 - R^2_H\right) \frac{N-1}{N-K-1}$, with N = the sample size and K the number of surrogates. The `Corr.R2_H` component contains the R^2_H that use the latter correction.
- `Lower.Dig.Corrs.Sigma=` : A `data.frame` that contains the matrix of the identifiable and unidentifiable correlations in $\boldsymbol{\Sigma}$ (lower diagonal elements) that were used to compute `R2_H` and `Corr.R2_H` in each of the runs. This information is useful to e.g., identify the subset of R^2_H values that are in line with certain biologically plausible constraints. For example, one can select the subset of R^2_H for which $r(T_0, T_1) < 0$ based on this information.

- `Call=` : the function call.

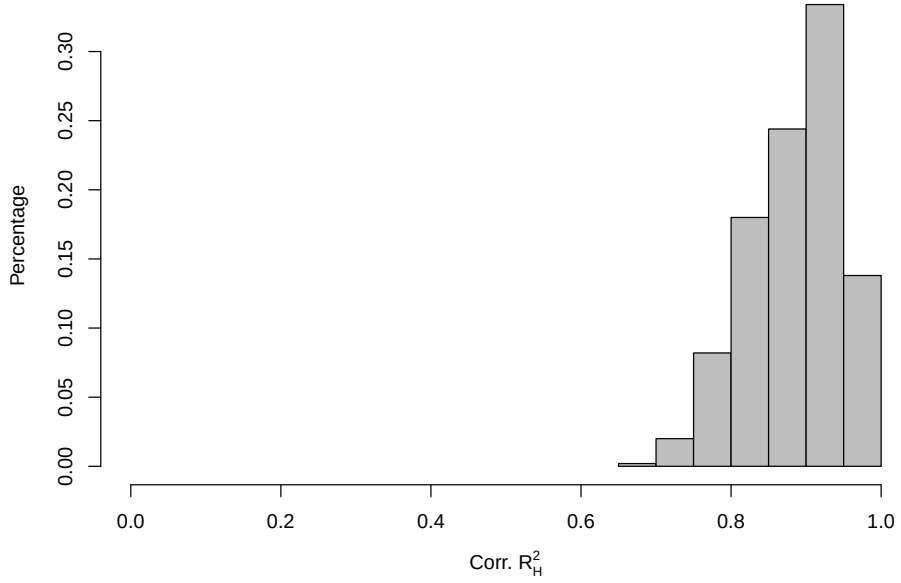
For example, the first 10 bias-corrected R^2_H values (contained in the `Corr.R2_H` component of the fitted `ICA_S1S2S3` object) can be obtained using the command:

```
> ICA_S1S2S3$Corr.R2_H[1:10]

# Generated output:
[1]  0.8290846 0.7874177 0.8902235 0.9099270 0.8971906
0.8983485 0.9026660 0.8296100 0.8840793 0.8849449
```

The results can be explored by applying the `summary()` and `plot()` functions to the fitted object:

```
> summary(ICA_S1S2S3)

# Generated output:

Function call:

ICA.ContCont.MultS(M = M, N = 15, Sigma = Sigma, Show.Progress = TRUE)



# Uncorrected R2_H results summary
#~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

Mean R2_H: 0.9097 (0.0481)  [min: 0.7508; max: 0.9986]
Mode R2_H:  0.9344

Quantiles of the distribution:

       5%       10%       20%       50%       80%       90%       95%
0.8240214 0.8402335 0.8666881 0.9182989 0.9521223 0.9672257 0.9812187



# Bias-corrected R2_H results summary
#~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

Mean adjusted R2_H: 0.8851 (0.0612)  [min: 0.6829; max: 0.9982]
Mode adjusted R2_H:  0.9165

Quantiles of the distribution:

       5%       10%       20%       50%       80%       90%       95%
0.7760273 0.7966608 0.8303303 0.8960168 0.9390648 0.9582873 0.9760965
```

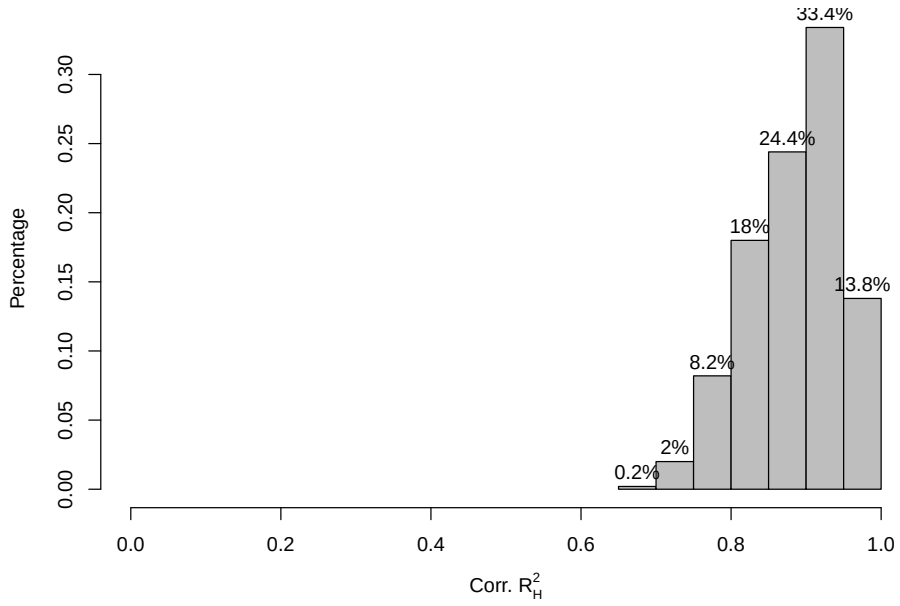
```
> plot(ICA_S1S2S3, xlim=c(0,1))

# Generated output:
```



The first part of the output of the `summary()` function shows descriptives of the R^2_H values. It can be seen that the mean, median, and mode of R^2_H are high with values 0.910, 0.918, and 0.934, respectively. Further, the impact of the unverifiable assumptions on the results is small, i.e., the range of R^2_H values is quite narrow and equals $[0.751, 0.999]$. The second part of the output of the `summary()` function provides the same descriptives for the bias-corrected R^2_H. As can be seen, the bias-corrected R^2_H values are slightly below those of R^2_H (as expected), but the results for both metrics are very similar and the qualitative conclusions are identical.

By default, the `plot()` function provides a histogram of the bias-corrected R^2_H values. A similar plot for the uncorrected R^2_H values can be obtained using the `R2_H=TRUE` argument in the `plot()` function call. Another useful plot option is `Labels=TRUE`, which adds the percentages of ICA values that fall within the different bins of the histogram. For example, a histogram of the R^2_H values with labels can be obtained using the following command:

```
> plot(ICA_S1S2S3, R2_H=TRUE, Labels = TRUE, ylim=c(0, 0.35))

# Generated output:
```



Considering a subset of the three surrogates In the analysis above, a multivariate analysis with three surrogates were used. It is straightforward to consider only one or two of the available surrogates in the analysis by respecifying $\mathbf{\Sigma}$. To conduct the analyses using all possible combinations of the available surrogates (i.e., S_1, S_2, S_3, $S_1 + S_2$, $S_1 + S_3$, $S_2 + S_3$, and $S_1 + S_2 + S_3$), the following commands can be used:

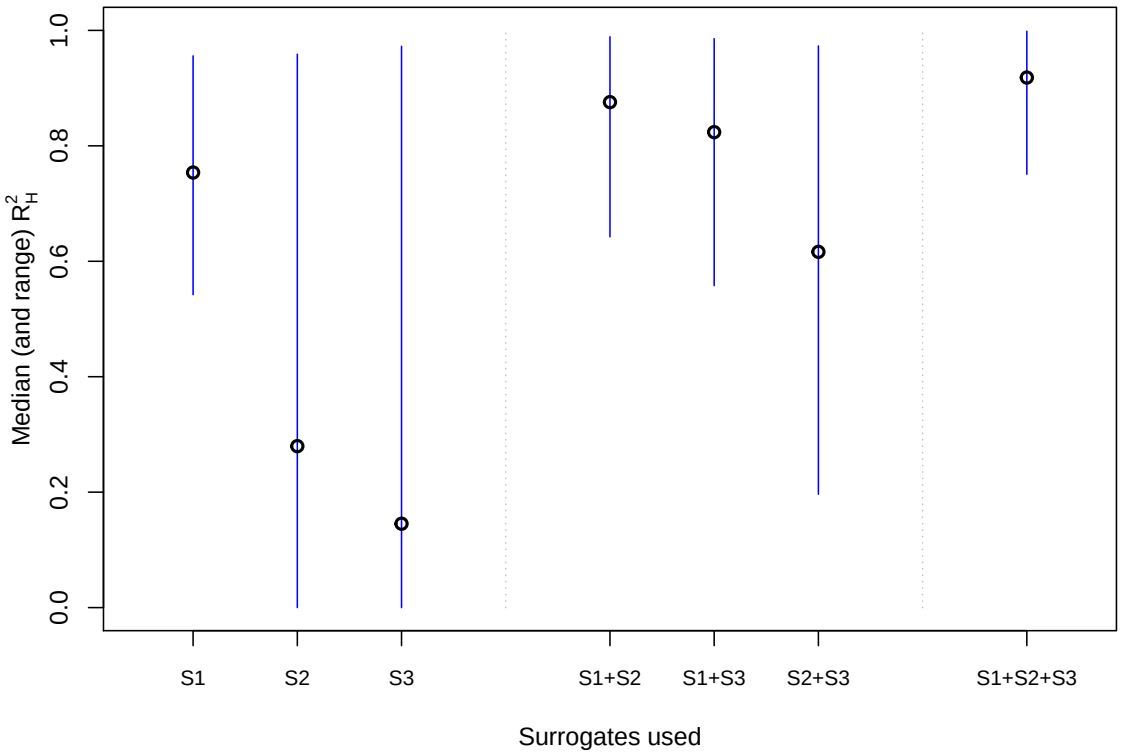
```
 # S1 alone
> Sigma_T_S1 <- Sigma[c(1:4), c(1:4)]
> ICA_S1 <- ICA.ContCont.MultS(Sigma = Sigma_T_S1, N=15)

# S2 alone
> Sigma_T_S2 <- Sigma[c(1,2,5,6), c(1,2,5,6)]
> ICA_S2 <- ICA.ContCont.MultS(Sigma = Sigma_T_S2, N=15)

# S3 alone
> Sigma_T_S3 <- Sigma[c(1,2,7,8), c(1,2,7,8)]
> ICA_S3 <- ICA.ContCont.MultS(Sigma = Sigma_T_S3, N=15)
```

```
# S1, S2
> Sigma_T_S1S2 <- Sigma[c(1,2,3,4,5,6), c(1,2,3,4,5,6)]
> ICA_S1S2 <- ICA.ContCont.MultS(Sigma = Sigma_T_S1S2, N=15)

# S1, S3
> Sigma_T_S1S3 <- Sigma[c(1,2,3,4,7,8), c(1,2,3,4,7,8)]
> ICA_S1S3 <- ICA.ContCont.MultS(Sigma = Sigma_T_S1S3, N=15)

# S2, S3
> Sigma_T_S2S3 <- Sigma[c(1,2,5,6,7,8), c(1,2,5,6,7,8)]
> ICA_S2S3 <- ICA.ContCont.MultS(Sigma = Sigma_T_S2S3, N=15)

# S1, S2, S3
> ICA_S1S2S3 <- ICA.ContCont.MultS(Sigma = Sigma, N=15)
```

The results of these analyses can be summarized in a plot that shows the medians (small circles) and the ranges (blue lines) of the R^2_H values in the different scenarios:

```
# Make empty plot field
> plot(x=c(.5:9.5), y=rep(0, times=10), col=0, ylab=" ", xaxt="no",
+      xlab="Surrogates used", ylim=c(0, 1))
> mtext(expression(paste("Median (and range) ", R[H]^2)), side=2, line = 2)

# Add results for S1
> segments(x0 = 1, x1 = 1, y0 = min(ICA_S1$R2_H),
+   y1 = max(ICA_S1$R2_H),
+   col="blue", lwd=1); points(x=c(1), y=median(ICA_S1$R2_H), lwd=2)
# Add results for S2
> segments(x0 = 2, x1 = 2, y0 = min(ICA_S2$R2_H),
+   y1 = max(ICA_S2$R2_H),
+   col="blue", lwd=1); points(x=c(2), y=median(ICA_S2$R2_H), lwd=2)
# Add results for S3
> segments(x0 = 3, x1 = 3, y0 = min(ICA_S3$R2_H),
+   y1 = max(ICA_S3$R2_H),
+   col="blue", lwd=1); points(x=c(3), y=median(ICA_S3$R2_H), lwd=2)

# Add verticle grey line
> segments(x0 = 4, x1 = 4, y0=0, y1=1, col="grey", lty=3)

# Add results for S1 + S2
> segments(x0 = 5, x1 = 5, y0 = min(ICA_S1S2$R2_H),
+   y1 = max(ICA_S1S2$R2_H),
+   col="blue", lwd=1); points(x=c(5), y=median(ICA_S1S2$R2_H), lwd=2)
# Add results for S1 + S3
```

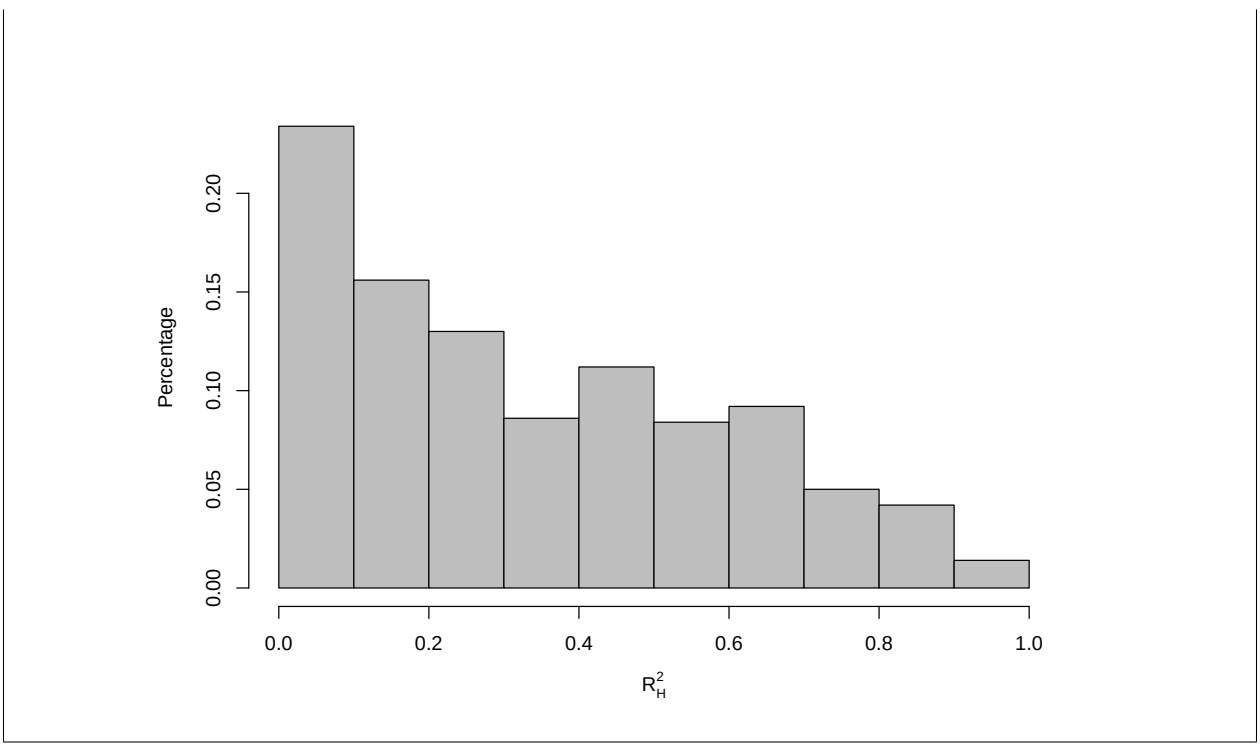
```
> segments(x0 = 6, x1 = 6, y0 = min(ICA_S1S3$R2_H),
+   y1 = max(ICA_S1S3$R2_H),
+   col="blue", lwd=1); points(x=c(6), y=median(ICA_S1S3$R2_H), lwd=2)
# Add results for S2 + S3
> segments(x0 = 7, x1 = 7, y0 = min(ICA_S2S3$R2_H),
+   y1 = max(ICA_S2S3$R2_H),
+   col="blue", lwd=1); points(x=c(7), y=median(ICA_S2S3$R2_H), lwd=2)

# Add verticle grey line
> segments(x0 = 8, x1 = 8, y0=0, y1=1, col="grey", lty=3)

# Add results for S1 + S2 + S3
> segments(x0 = 9, x1 = 9, y0 = min(ICA_S1S2S3$R2_H),
+   y1 = max(ICA_S1S2S3$R2_H),
+   col="blue", lwd=1); points(x=c(9), y=median(ICA_S1S2S3$R2_H), lwd=2)
# Add labels X-axis
> axis(1, at=c(1, 2, 3, 5, 6, 7, 9), cex.axis=.8,
+   labels = c("S1", "S2", "S3", "S1+S2", "S1+S3", "S2+S3", "S1+S2+S3"))

# Generated output:
```



As can be seen, the ranges of the estimated R^2_H values were very wide in the univariate setting, and their median R^2_H relatively low. For example, in the univariate analysis in which S_2 was used as a surrogate, the median R^2_H equalled 0.280 and its range was $[0.001, 0.959]$. These results can be obtained by using the following command:

```
> summary(ICA_S2)

# Generated output:

Function call:

ICA.ContCont.MultS(N = 15, Sigma = Sigma_T_S2)



# Uncorrected R2_H results summary
```

```
#~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

Mean R2_H: 0.3369 (0.2582)  [min: 0.0000; max: 0.9587]
Mode R2_H:  0.0812

Quantiles of the distribution:

         5%         10%         20%         50%         80%         90%         95%
0.007002998 0.022066935 0.073900620 0.279684511 0.598672586 0.703471548 0.805695627



# Bias-corrected R2_H results summary
#~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

Mean adjusted R2_H: 0.2859 (0.2780)  [min: 0.0000; max: 0.9555]
Mode adjusted R2_H:  0.0105

Quantiles of the distribution:

          5%          10%          20%          50%          80%          90%
-0.069381386 -0.053158685  0.002662206  0.224275627  0.567801247  0.680661667
         95%
 0.790749137


# Request plot of R2H
> plot(ICA_S2, xlim=c(0,1), R2_H = TRUE, Corr.R2_H = FALSE)

# Generated output:
```



The best prediction of ΔT was obtained when all three surrogates were used, with median $R^2_H = 0.918$ and range $[0.751, 0.999]$ (see output shown above). Notice that even though S_2 and S_3 had limited value as univariate surrogates, the use of these two surrogates led to a substantial increase in R^2_H when *considered jointly* with S_1. This shows that S_2 and S_3 contain useful information that is not captured by S_1, and it supports the claim that the use of multivariate surrogates can be a route to ameliorate the problems of finding a good surrogate.

3 The multivariate adjusted association

The function `AA.MultS` in the *Surrogate* package can be used to compute the Adjusted Association (AA) in the Multiple-Surrogate setting. The multivariate adjusted association γ^2_Δ is defined as the treatment-corrected squared multiple correlation between T and its best linear predictor based on the treatment-corrected S_1, S_2, ..., and S_k (for details, see Van der Elst *et al.*, 2018). The `AA.MultS` function requires the following arguments:

- `Sigma_gamma=` : the variance-covariance matrix of the residuals of regression models in which the true and surrogate endpoints are regressed on the treatment, referred to as $\mathbf{\Sigma}_\gamma$.
- `N=` : the number of patients in the dataset.

For the transPAT case study, the variance-covariance matrix of the residuals is:

$$\boldsymbol{\Sigma}_{\gamma} = \begin{pmatrix} 93.4608449102 & 0.0467089011 & -0.9736034172 & -0.0007668859 \\ 0.0467089011 & 0.0000311512 & -0.0003008816 & 0.0000000116 \\ -0.9736034172 & -0.0003008816 & 0.0327815216 & 0.0000857586 \\ -0.0007668859 & 0.0000000116 & 0.0000857586 & 0.0000003877 \end{pmatrix}.$$

Using the `AA.MultS` function, the multivariate adjusted association in which S_1, S_2 and S_3 are considered can now be computed with the command:

```
# First define the Sigma_gamma matrix
> Sigma_gamma <- matrix(data = c(
93.4608449102,  4.670890e-02, -9.736034e-01, -7.668859e-04,
0.0467089011,  3.115116e-05, -3.008816e-04,  1.163247e-08,
-0.9736034172, -3.008816e-04,  3.278152e-02,  8.575864e-05,
-0.0007668859,  1.163247e-08,  8.575864e-05,  3.877270e-07
), nrow=4)

# Conduct the analysis
> AA_S1S2S3 <- AA.MultS(Sigma_gamma = Sigma_gamma, N = 15, Alpha = .05)
```

The results can be examined by applying the `summary()` function to the fitted `AA_S1S2S3` object:

```
> summary(AA_S1S2S3)

# Generated output

Function call:

AA.MultS(Sigma_gamma = Sigma_gamma, N = 15, Alpha = 0.05)


# Uncorrected multivariate Adjusted Association
# Fisher Z 95%-based confidence interval (N = 15)
#~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

  Multivariate AA Standard Error CI lower limit CI upper limit
           0.8862         0.0618         0.7650         1.0000


# Bias-corrected multivariate Adjusted Association
# Fisher Z 95%-based confidence interval (N = 15)
#~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

  Adjusted multivariate AA Standard Error CI lower limit CI upper limit
                    0.8552         0.0773         0.7036         1.0000
```

The first part of the output of the summary function provides the estimated multivariate adjusted association and its 95% confidence interval, i.e., 0.886 [0.765, 1.000]. As was also the case with R^2_H, the second part

of the output provides the bias-corrected multivariate adjusted association, which was similar and equalled 0.885 [0.704, 1.000].

Considering a subset of the available surrogates In the above analysis, all three surrogates were used. It is straightforward to consider less than three surrogates by simply respecifying the appropriate $\mathbf{\Sigma}_\gamma$. For example, γ_Δ^2 using S_1 and S_3 as surrogates can be obtained using the commands:

```
# Conduct the analysis
> Result <- AA.MultS(Sigma_gamma = Sigma_gamma[c(1,2,4), c(1,2,4)],
N = 15, Alpha = .05)



# Generated output:
> summary(Result)

Function call:

AA.MultS(Sigma_gamma = Sigma_gamma[c(1, 2, 4), c(1, 2, 4)], N = 15,
    Alpha = 0.05)


# Uncorrected multivariate Adjusted Association
# Fisher Z 95%-based confidence interval (N = 15)
#~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

  Multivariate AA Standard Error CI lower limit CI upper limit
           0.7663         0.1181         0.5349         0.9978


# Bias-corrected multivariate Adjusted Association
# Fisher Z 95%-based confidence interval (N = 15)
#~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

  Adjusted multivariate AA Standard Error CI lower limit CI upper limit
                    0.7274         0.1342         0.4643         0.9905
```

4 Conclusion

In conclusion, the univariate analyses based on R_H^2 and γ_Δ^2 indicated that no 'good' surrogate could be established for T. However, when all three surrogates were considered in the analysis, the R_H^2 and γ_Δ^2 values were high. These results fully support the claim that the use of multivariate surrogates can be a useful strategy to identify 'good' surrogates.

Part III

An additional simulation study.

To further explore the impact of adding an extra surrogate on the obtained R^2_H, an additional simulation study was conducted where the focus is on bivariate setting $\boldsymbol{S} = (S_1, S_2)$. A total of four scenarios were considered. These are detailed below.

1. Scenario 1: all identifiable correlations between the surrogates and the true endpoint, and between the surrogates themselves are low and equal 0.3. Assumed correlation structure:

	T_0	T_1	S_{10}	S_{12}	S_{20}	S_{21}
T_0	1.0					
T_1	-	1.0				
S_{10}	0.3	-	1.0			
S_{11}	-	0.3	-	1.0		
S_{20}	0.3	-	0.3	-	1.0	
S_{21}	-	0.3	-	0.3	-	1.0

   Note. '-' refers to unidentifiable correlations for which no restrictions are set.

2. Scenario 2: same as scenario 1, but the identifiable correlations between S_1 and T in both treatment conditions now equal 0.8. Assumed correlation structure:

	T_0	T_1	S_{10}	S_{12}	S_{20}	S_{21}
T_0	1.0					
T_1	-	1.0				
S_{10}	0.8	-	1.0			
S_{11}	-	0.8	-	1.0		
S_{20}	0.3	-	0.3	-	1.0	
S_{21}	-	0.3	-	0.3	-	1.0

3. Scenario 3: same as scenario 1, but the identifiable correlations between S_2 and T in both treatment conditions now equal 0.8. Assumed correlation structure:

	T_0	T_1	S_{10}	S_{12}	S_{20}	S_{21}
T_0	1.0					
T_1	-	1.0				
S_{10}	0.3	-	1.0			
S_{11}	-	0.3	-	1.0		
S_{20}	0.8	-	0.3	-	1.0	
S_{21}	-	0.8	-	0.3	-	1.0

4. Scenario 4: same as scenario 1, but the identifiable correlations between S_1 and T and between S_2 and T in both treatment conditions now equal 0.8. Assumed correlation structure:

	T_0	T_1	S_{10}	S_{12}	S_{20}	S_{21}
T_0	1.0					
T_1	-	1.0				
S_{10}	0.8	-	1.0			
S_{11}	-	0.8	-	1.0		
S_{20}	0.8	-	0.3	-	1.0	
S_{21}	-	0.8	-	0.3	-	1.0

In the simulations, the impact of adding additional surrogates on R^2_H (in particular its range) was examined. R^2_H was estimated using the sensitivity-based approach that was proposed in Van der Elst *et al.* (2018). The function `ICA.ContCont.MultS()` of the *Surrogate* package (for details, see Section 2 of Part II of this Web Appendix) was used to conduct the analysis with $M = 500$ runs of the algorithm.

Results Figure 1 shows histograms of R^2_H using S_1 alone (top panels) and using both S_1 and S_2 (bottom panels) in all four scenarios. Table 1 shows summary statistics of R^2_H in the different analyses.

In scenario 1, all identifiable correlations were low. As expected, most of the R^2_H values were low as well and ICA spanned nearly the entire unit interval in both the univariate and bivariate analyses.

In scenario 2, S_1 had a relatively strongly correlation with T (i.e., $r = 0.80$) and S_2 had a low correlation with T (i.e., $r = 0.30$). The results show that most of the obtained R^2_H values are relatively high in the analysis where S_1 is used (see Figure 1, top panel, 2nd column), but the range of ICA is very large. When S_2 is added in the analysis, the R^2_H is not substantially affected (see Figure 1, top panel, 2nd column). This result indicates that the addition of a new candidate surrogate that is not strongly correlated with T does not lead to a substantial reduction of the range of R^2_H (i.e., no reduction in uncertainty).

In scenario 3, S_1 had a low correlation with T (i.e., $r = 0.30$) and S_2 had a high correlation with T (i.e., $r = 0.80$). Here the results show that most of the R^2_H values are low in the analysis where S_1 is used (see Figure 1, top panel, 3rd column) and ICA nearly spans the entire unit interval. Adding S_2 as an additional surrogate substantially increases R^2_H (see Figure 1, top panel, 3rd column), but the range of ICA remains large and spans nearly the entire unit interval.

Finally, in Scenario 4, both S_1 and S_2 were relatively highly correlated with T (i.e., $r = 0.80$). The results show that R^2_H is high when S_1 is considered alone, but the range of ICA is wide as well. Importantly, adding another candidate surrogate S_2 which is relatively highly correlated to T leads to a strong reduction in the range of R^2_H. Indeed, all R^2_H are now at least 0.75, and thus it can be concluded that $\mathbf{\Delta S}$ is a relatively good surrogate for ΔT in all realities that are compatible with the identifiable correlations.

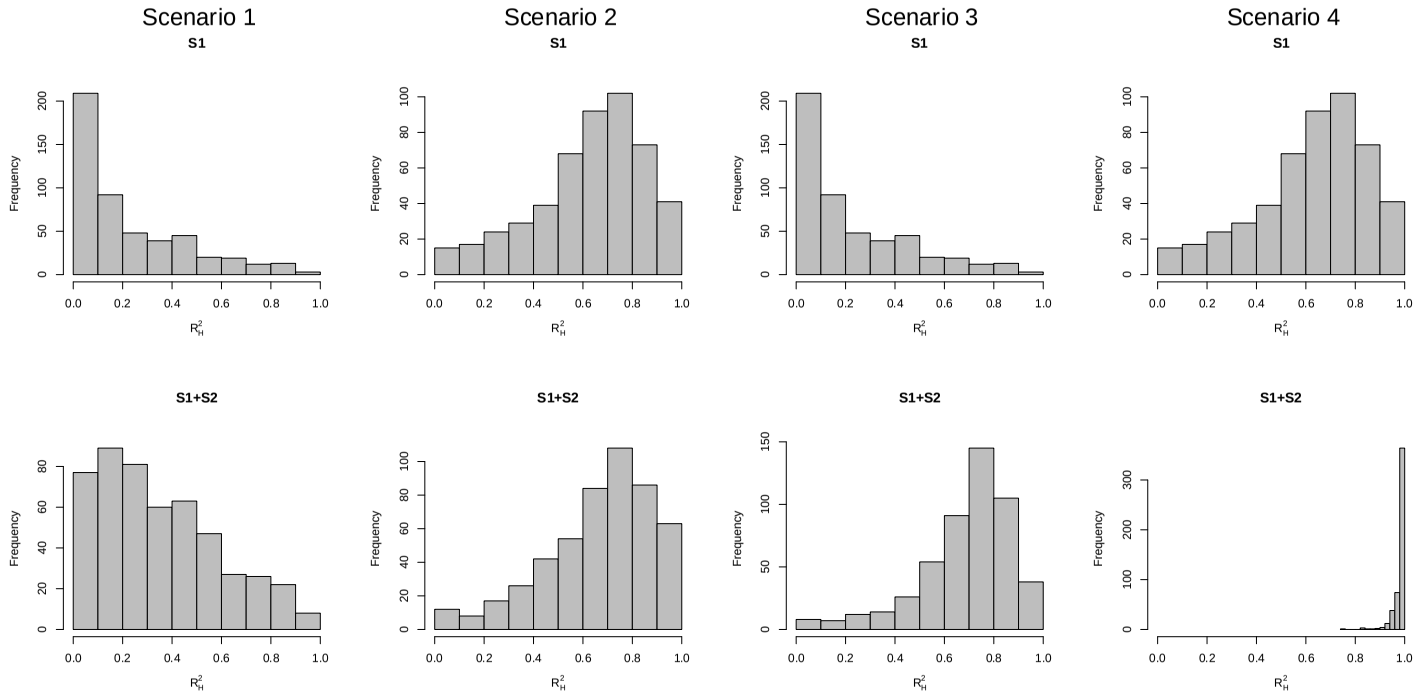


Figure 1: Simulation results. Upper panel: histograms of R^2_H using S_1 alone, bottom panel: histograms of R^2_H using both S_1 and S_2.

Table 1: Simulation study results. Descriptives of R^2_H for the analyses where only S_1 and both S_1 and S_2 were considered.

Scenario	Surrogates used	Mean	Median	Sd	Min	Max
1	S1	0.226	0.139	0.230	0.000	0.957
	S1+S2	0.349	0.301	0.240	0.000	0.994
2	S1	0.622	0.669	0.226	0.000	0.996
	S1+S2	0.665	0.705	0.216	0.013	0.997
3	S1	0.226	0.139	0.230	0.000	0.957
	S1+S2	0.688	0.726	0.185	0.027	0.986
4	S1	0.622	0.669	0.226	0.000	0.996
	S1+S2	0.981	0.989	0.024	0.750	1.000

References

Cohen, J., Cohen, P., West, S. G., and Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences. Third edition.* New Jersey: Lawrence Erlbaum Associates.

Van der Elst, W., Alonso, A. A. A., Geys, H., Bijnens, L., Meyvisch, P., Sengupta, R., and Molenberghs, G. (2018). Univariate versus multivariate surrogate endpoints in the single-trial setting.

Ruiz, V. E. *et al.* (2017). A single early-in-life macrolide course has lasting effects on murine microbial network topology and immunity. *Nature Communications* **8**, 1–14.

Yin, P., and Fan, X. (2001). Estimating R^2 shrinkage in multiple regression: A comparison of different analytical methods. *The Journal of Experimental Education, 69,* 203–224.