Methodology to validate surrogate endpoints in clinical trials: which level of evidence is fit for purpose?

ISPOR Workshop
November 5th, 2013

Dr. Hélène Karcher
Wilma Erhardt
Dr. Carsten Schwenke
Theophile Bigirumurame

Outline

Context

- A new era for surrogates
- Case examples of surrogate approvals and rejections

Practical surrogacy validation

- Methods and their limitations
- New regulations *explained* (IQWiG & NICE)
- Working example of model-based surrogacy validation

Context

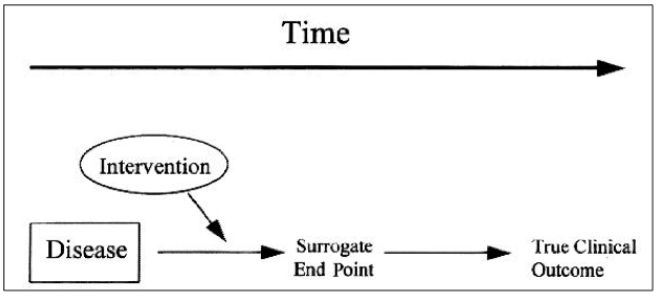
- A new era for surrogates
- Case examples of surrogate approvals and rejections

Definition : surrogate endpoint

- A physical sign or laboratory measurement to be used instead of a clinical endpoint.
  - To evaluate the impact of an intervention on the pathological process.
  - Can be measured earlier, more conveniently, more frequently and/or less invasively than the clinically-significant endpoint.
  - Correlated to the clinically-significant endpoint in individual patients.
  - Treatment impact on surrogate predicts treatment impact on the clinically-significant endpoint.

Perfect situation

- One causal route from disease to outcome
- Surrogate is an intermediate step in the route
- Treatment has no impact on relationship of surrogate and outcome



(Fleming & DeMets, Annals of Internal Medicine 1996, 125: 605-613)

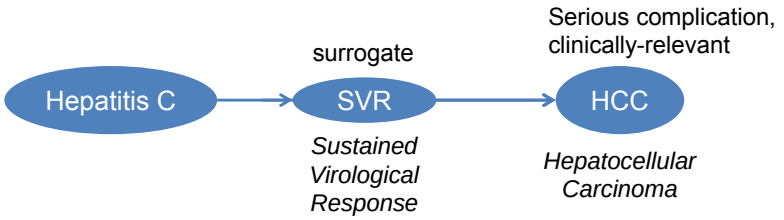
Questions to the audience

- Examples of a surrogate and the endpoint it is replacing?
- Which of these surrogates are ‚valid'? not valid? formally validated?
- For surrogates that are „valid", why?
- What consistutes an „adequate" surrogate validation?

Examples of routinely-used surrogates

Disease	Surrogate	Clinical Outcome
Osteoporosis	Bone density	Fractures
Cardiovascular diseases	Blood pressure	MI, OS
HIV	CD-4 cell count	OS
Diabetes mellitus	HbA1c in blood	Complications, OS, QoL
Hepatitis C	SVR - Sustained virologic response	Remission HCC: hepatocellular carcinoma

Surrogate validity/validation drives strength of the evidence



- Boceprevir and Telaprevir assessments
  - SVR is valid, but not formally validated
  - No formal validation possible for ethical reasons

Consequence: IQWiG downgraded the products' additional benefits to „not quantifiable"

Questions to the audience

- Example where a surrogate did not correctly predict the endpoint of interest?
- Why was the surrogate thought to be valid?

Where surrogates lead wrong

Cause	Recommendation	Surrogate	Outcome
Cardiovascular disease	Decrease of lipids	Serum cholesterol ↓ Blood glucose level ↓	Incidence CAD ↓ OS ↑
Osteoporosis	Natrium-Fluorid	Bone density ↑	Incidence fractures ↑

Growing need for surrogates

- Treatments become more effective. Endpoint is reached later.
  - long follow-up needed to observe the outcome
  - small effect size for outcome when observed early

  surrogates allow to read out clinical trial results earlier
- Ethical issue in stopping a trial early once the investigational product is proven to be superior based on other outcomes

  infer results on clinically-significant endpoints in cases where appropriate trials for these results are not ethically feasible

  Unbiased clinical endpoint cannot be observed due to :
  - Cross-over to efficacious therapy
  - Emergence of many possible subsequent therapies

Approval and reimbursement decisions often need to be based on surrogates

Case example : bevacizumab (Avastin)

Context

- Indication : chemo-naïve metastatic HER2-negative breast cancer
- Clinical endpoint : overall survival (OS)
- Surrogate endpoint : progression-free survival (PFS)

Approval process

- 2008 FDA approved Avastin based on PFS improvement
  - No data on OS nor quality-of-life improvement
- 2011 FDA finally revoked their approval
- EMA continued to approve Avastin for breast cancer

Case example : renal cell carcinoma

Context

- Indication : metastatic renal cell carcinoma (mRCC)
- Clinical endpoint : overall survival (OS)
- Surrogate endpoint : progression-free survival (PFS)

Approval process

- All agents approved based on PFS improvement only, no OS data (with 2 exceptions)
- After approval, retrospective studies* showed improvement of OS in real life (sequence of therapies)

* e.g., Kamba et al. 2013. Improvement of prognosis in patients with metastatic renal cell carcinoma […]. Int J Clin Oncol.

Case example : vandetanib

Context

- Indication : metastatic medullary thyroid carcinoma
- Clinical endpoint : overall survival (OS)
- Surrogate endpoint : progression-free survival (PFS)

Approval process

- Conditional approval by EMA with orphan drug status
- 2012: G-BA decided no additional benefit
- 2013: G-BA decided minor additional benefit based on pain progression and PFS (PFS no surrogate for OS but individual morbidity endpoint)

Characteristics of successfully developed and accepted surrogates[1,2]

- Biologic plausibility

- Accurate prognosis of disease outcome
- Accurate association between a change in the surrogate endpoint caused by intervention, and ultimate disease outcome

1. Buyse M, Sargent DJ, Grothey A, Matheson A, de Gramont A. Biomarkers and surrogate end points--the challenge of statistical validation. Nature reviews Clinical oncology. 2010;7(6):309-17.
2. Lathia CD, Amakye D, Dai W, Girman C, Madani S, Mayne J, MacCarthy P, Pertel P, Seman L, Stoch A, Tarantino P, Webster C, Williams S, Wagner JA. The value, qualification, and regulatory use of surrogate end points in drug development. Clinical pharmacology and therapeutics. 2009;86(1):32-43.

Validation methods and their limitations

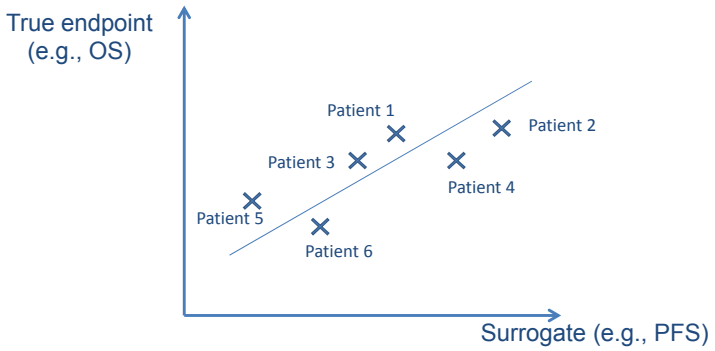
Method to validate a surrogate

- Biologic plausibility
  1. Conduct literature reviews
- Accurate prognosis of disease outcome
  2. Demonstrate 'Individual-level surrogacy'
- Accurate association between a change in the surrogate endpoint caused by intervention, and ultimate disease outcome
  3. Demonstrate 'Trial-level surrogacy'

Literature review on causality

- Systematic review to gather evidence for the causal relationship disease - surrogate - endpoint
  - Biological and clinical arguments
  - predictive and pronostic factors for surrogate and true endpoint

Individual-level surrogacy validation

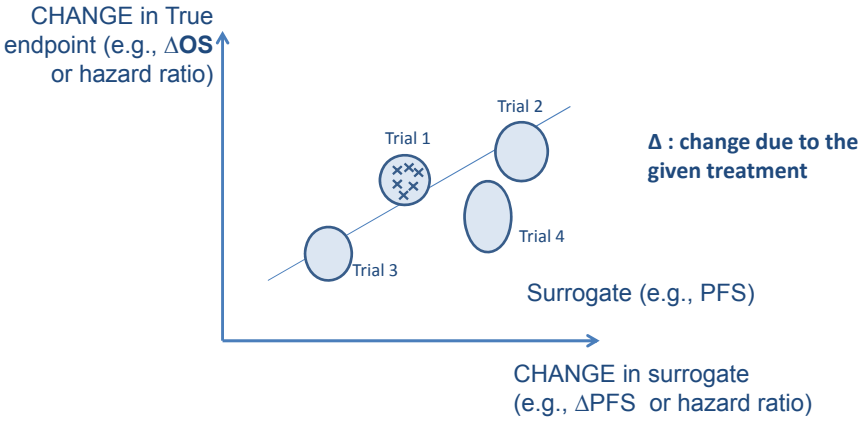


→ Tests correlation between surrogate and endpoint for a particular treatment, over a restricted range. Part 2 of surrogacy validation.

Limitations

- Requires individual patient data from at least one trial

Trial-level surrogacy validation



→ Tests correlation between ΔPFS and ΔOS for several treatments. Part 3 of surrogacy validation.

Limitations

- Requires trial-level data from multiple trials or units
  - Trials in the same indication
  - Trials with drugs or interventions of the same class
    - Similar action on the disease to clinical endpoint pathway

HTA agencies on surrogate validation

HTA agencies & surrogacy validation

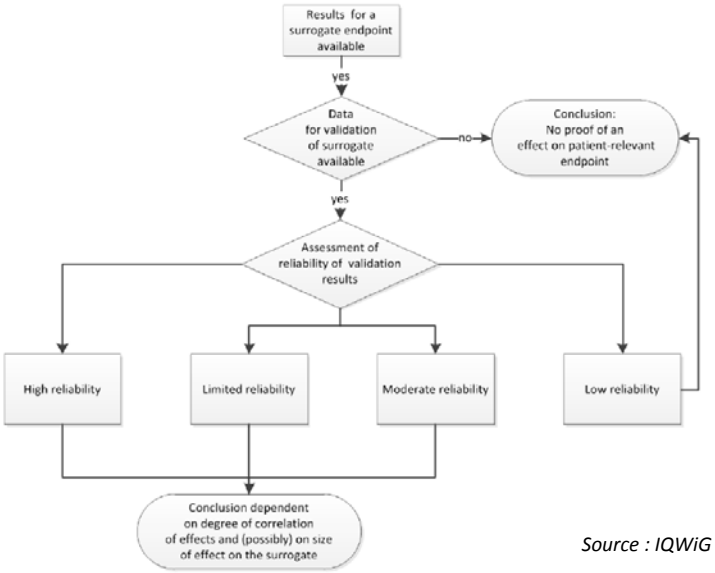
> 2011. IQWiG. [Validity of surrogate parameters in oncology] (Rapid report). Aussagekraft von Surrogatendpunkten in der Onkologie.

> 2012. NICE. Review of the guide to the methods of technology appraisal.

Surrogate – IQWiG requirements

- Measure of correlation (R^2) und Surrogate threshold effect (STE)
- Proof of correlation between surrogate and outcome *not sufficient*
- Validitation specific to the disease and indication
- Basis for evaluation: preferably meta-analyses
- Confidence intervals of correlation should be considered, especially in case of low correlations (STE)

IQWiG's benefit risk assessment depends on the reliability of surrogacy validation



Source : IQWiG

Surrogate evalution according to IQWiG's ‚Rapid report' *

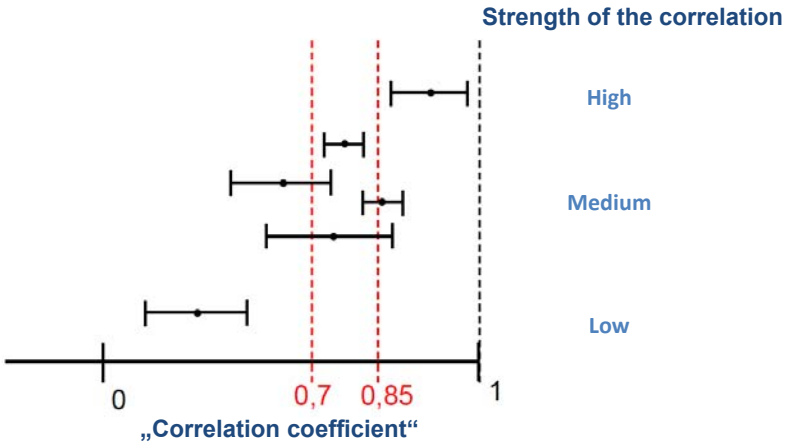
- Appropriate evaluation of the validity of evidence
- Evaluation is based on correlation of surrogate and outcome and confidence in evidence

Confidence in evidence	Correlation	Validity
High	High	Yes
	Moderate	Unclear
	Low	No
Moderate		unclear
Low		
Very low		

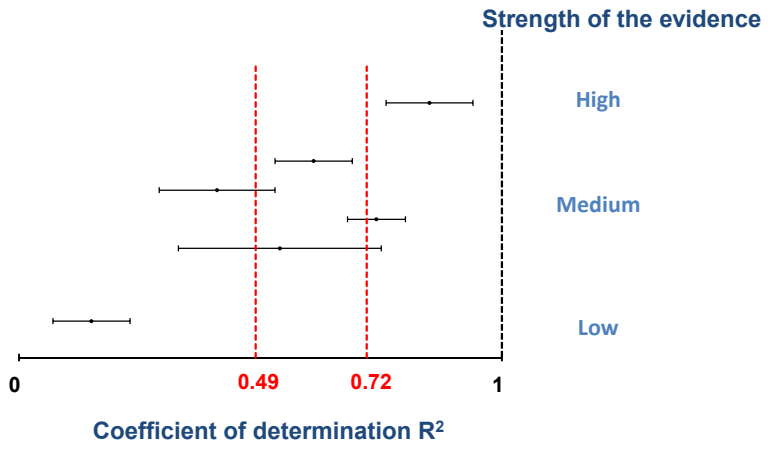
- Effect size
- Assessment of futher information on outcome from studies

* Aussagekraft von Surrogatendpunkten in der Onkologie. Rapid Report IQWiG-Berichte - Jahr: 2011 Nr. 80

IQWiG's position to evaluate trial and individual correlation



IQWiG's position in terms of coefficient of determination



Confidence in Evidence*

- High confidence in evidence, at least moderate correlation
  - Proof (could be downgraded to hint)
  - hint (80% CI of effect for surrogate > STE)
- Moderate confidence in evidence, at least moderate correlation
  - Hint (could be downgraded to clue)
  - Clue (80% CI of effect for surrogate > STE)
- Low confidence in evidence, at least moderate correlation
  - High correlation: clue (could be downgraded to none)
  - Moderate correlation: clue (95% KI of effect for surrogate > STE)

* Rapid Report A10-05

NICE requirements for validation

Clinical endpoints that reflect how a patient feels, functions, or how long a patient survives are regarded as more informative than surrogate endpoints (such as laboratory tests and imaging findings). When the use of 'final' clinical end points is not possible and 'surrogate' data on other outcomes is used to infer the effect of treatment on mortality and health-related quality of life, validation of the surrogate-to-final endpoint outcome relationship must be provided, together with details of the information on which the relationship is based and an explanation of how this relationship is quantified for use in modelling. The usefulness of the surrogate endpoint for estimating QALYs will be greatest when there is strong evidence that it predicts health-related quality of life and/or survival. In all cases, the uncertainty associated with the relationship between the endpoint and health-related quality of life or survival should be explored and quantified.

Source : NICE. 2012. Review of the guide to the methods of technology appraisal

universiteit hasselt | KU LEUVEN | I-BioStat

janssen

METHODOLOGY TO VALIDATE SURROGATE ENDPOINTS IN CLINICAL TRIALS: A CASE STUDY

Theophile Bigirumurame

Interuniversity Institute for Biostatistics and statistical Bioinformatics
Hasselt University, Belgium

Joint work with

Ziv Shkedy, Tomasz Burzykowski,

Hasselt University ,Belgium

Joris Diels, Suzy Van Sanden, Pushpike Thilakarathne,

Janssen, Beerse, Belgium

Case study

- Randomized, multicenter study in oncology.
- Each center is a different country.
- New treatment vs control.
- Primary endpoint: overall survival.
- Question: does treatment improve overall survival?

Can we use progression free survival as a surrogate?

Case Study

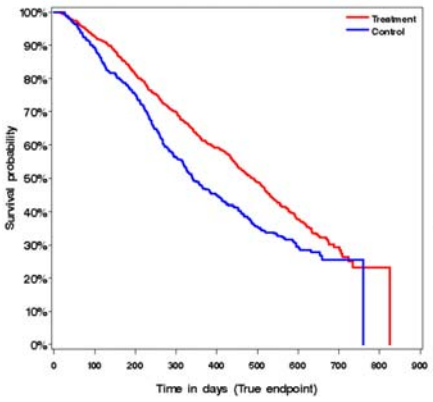
- Multicenters clinical trial conducted in 13 centers.

	MEDIAN SURVIVAL TIME (days)	
ARM / ENDPOINT	TREATMENT	CONTROL
Overall survival (True endpoint)	485	340
Progression free survival (Surrogate endpoint)	88	85

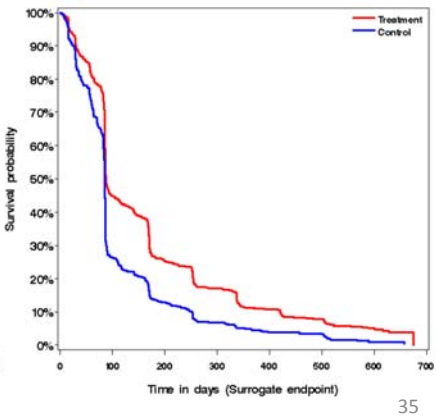
Case study

- Significant difference between treatment and control (P-value <0.002 and <0.001).

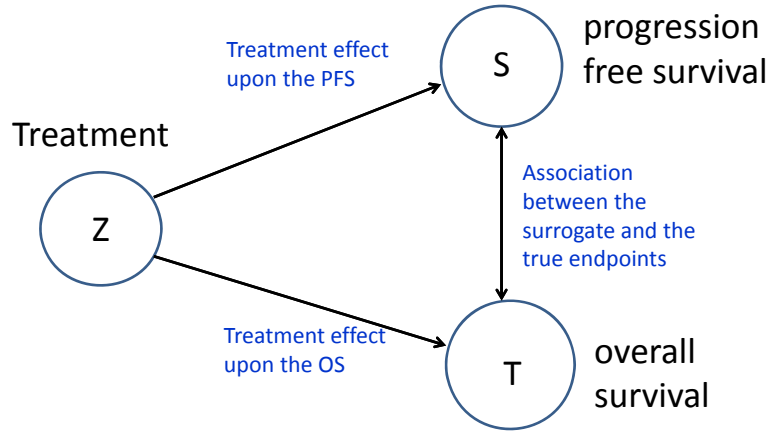
Overall survival



Progression free survival



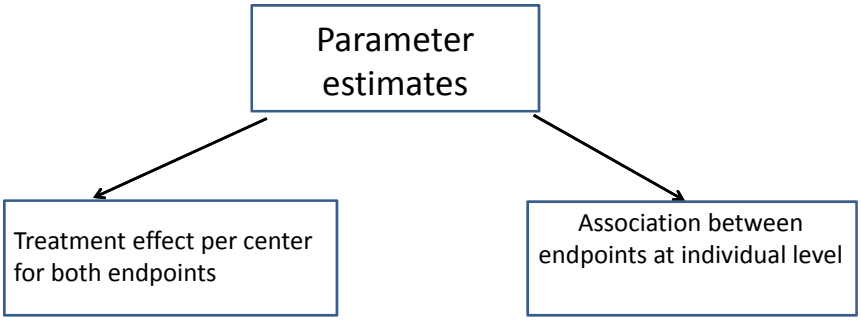
The surrogacy setting:



Surrogacy question:
can we use treatment effect on PFS to predict treatment effect on OS?

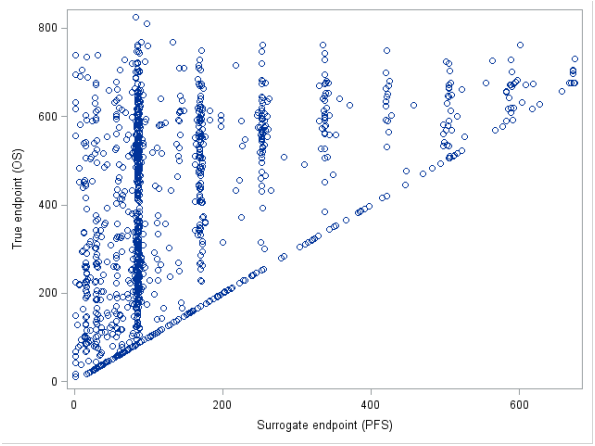
Methodology

- Jointly model the survival endpoints.
- OS=f(treatment,center).
- PFS=f(treatment,center).



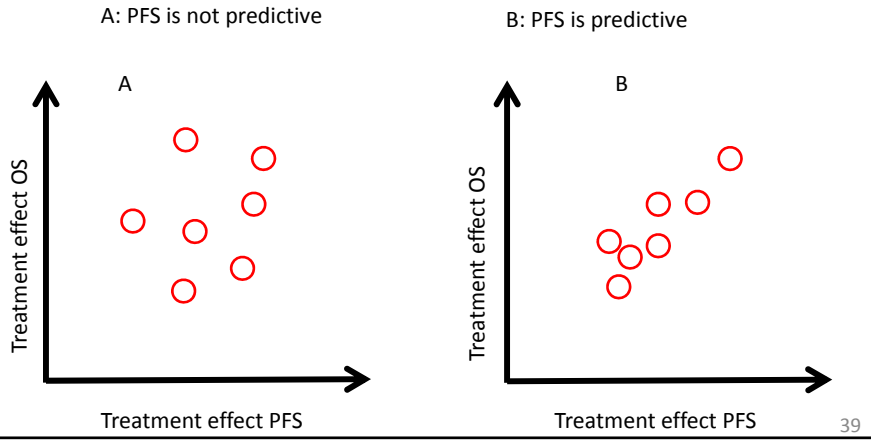
Surrogacy Question: individual level surrogacy

Is progression free survival associated on overall survival
After correcting for treatment?



Surrogacy Question: trial level surrogacy

Can we use treatment effect on progression free survival to predict treatment effect on overall survival?



Assessment of surrogacy levels

- Individual level surrogacy measured by Kendal's tau coefficient.
- Trial level surrogacy can be measured using the correlation between the treatment effects (R^2).

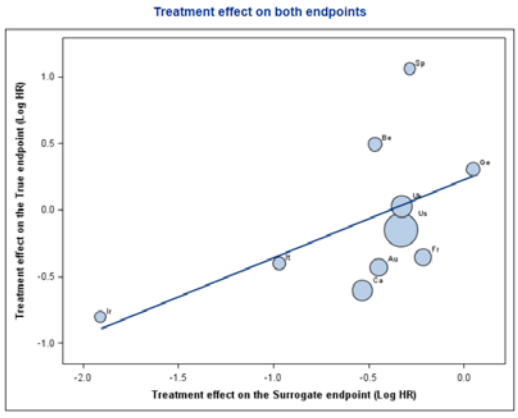
$$\hat{\beta}_i = \lambda_0 + \lambda_1 \hat{\alpha}_i + \varepsilon_i$$

- Treatment effect on the true endpoint
- Treatment effect on the surrogate endpoint
- Error terms

Assessment of surrogacy levels

- Good surrogate is expected to have R^2 close to 1 , at both level of surrogacy.
- Confidence interval around R^2 should be narrow.

Results



- Trial level R^2
  0.32 (0;0.79).
- Individual level
  Kendal's Tau
  0.30 (0.26;0.34).

SAS macros for other type of endpoints

True endpoint	Surrogate endpoint	Method
Survival	Survival	Fixed effect model
Survival	Survival	Two stage approach
Survival	Survival	Joint model (copula)
Survival	Binary	Joint model (fixed effect)
Normal	Normal	Joint model (fixed effect model)
Normal	Normal	Joint model (random effect model)
Normal	Binary	Joint model (Fixed effect model)

Discussion and conclusion of the case study

- Relatively low value of correlation between the treatment effect of PFS to the treatment effects of OS.
- Relative low value for the Kendal's tau coefficient.
- Based on the data used here, progression free survival does not seem to be a good surrogate for the overall survival.
- Software will be available online soon !!!

Thank you for your attention

helene.karcher@la-ser.com
werhardt@its.jnj.com
carsten.schwenke@scossis.de
theophile.bigirumurame@uhasselt.be

Example for the SAS macro

Software: SAS macro

%**Surv**(data=,surr=,surrind=,true=,trueind=, trt=,center=)

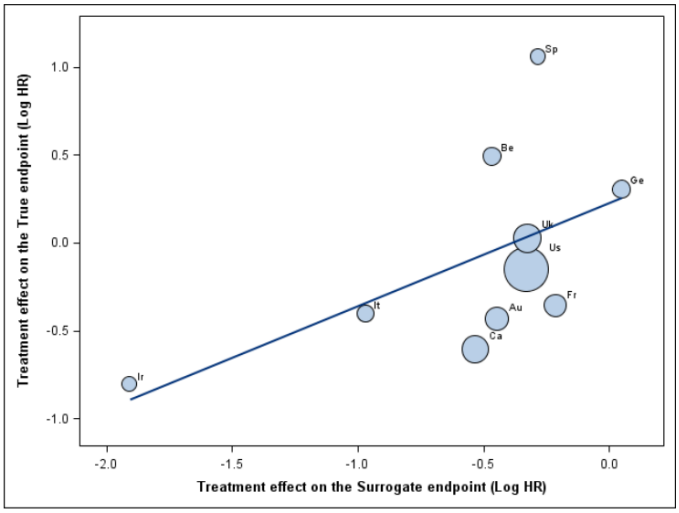
ID	TRUE	TRUEIND	SURR	SURRIND	TRT	CENTER
1	451	1	85	1	1	1
2	524	1	169	1	1	1
3	419	1	85	1	0	1
4	213	1	213	1	0	2
5	509	0	88	1	1	2
6	516	1	516	1	1	3
7	117	1	117	1	0	3
8	178	1	144	1	1	4
9	317	1	86	1	1	4
10	403	1	56	1	0	4
11	483	1	85	1	1	5
12	62	1	62	1	1	5
13	559	0	353	1	0	5

Output

VALIDATION USING CLAYTON COPULA

Survival Curve - Surrogacy measures - Effects plot -

Treatment effect on both endpoints



Output

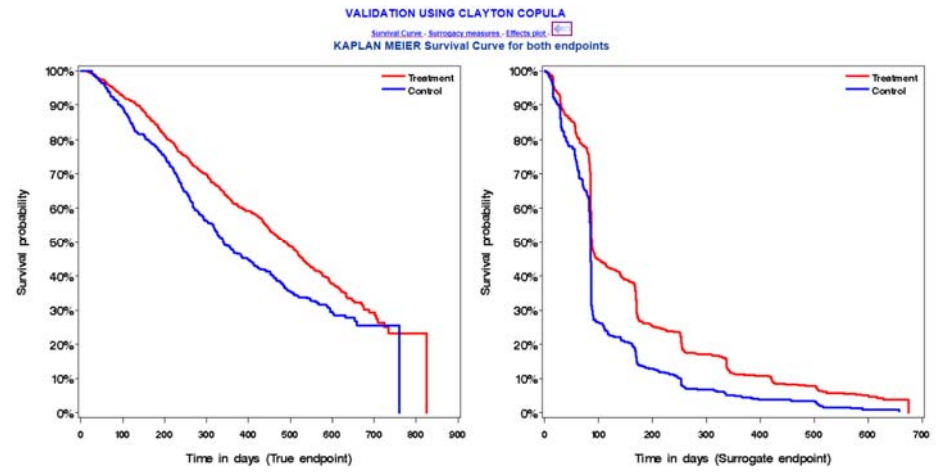
VALIDATION USING CLAYTON COPULA

Survival Curve - Surrogacy measures - Effects plot -

SURROGACY MEASURES

INDIVIDUAL			TRIAL		
LOWER	TAU	UPPER	LOWER	R square	UPPER
0.2635	0.3039	0.3442	-0.1592	0.3175	0.7942

Output



References

- Burzykowski, T., Molenberghs, G., Buyse, M., Renard, D., and Geys, H. (2001) Validation of surrogate endpoints in multiple randomized clinical trials with failure-time endpoints. *Applied Statistics*, **50**, 405-422.
- Burzykowski, T., Molenberghs,G., Buyse, M (2005). The evaluation of surrogate endpoints, Springer series.