

# Validität von Surrogatparametern

## 3. Gesundheitsforum zur Nutzenbewertung im Gesundheitswesen

Gesundheitsforschungsrat  
Institut für Wirtschaftlichkeit im Gesundheitswesen  
IQWiG

Prof. Heiner C. Bucher MD MPH  
Basel Institute for Clinical Epidemiology & Biostatistics  
University Hospital Basel  
[www.ceb-institute.org](http://www.ceb-institute.org)

# Overview

Definition of a surrogate endpoint

Criteria for the use of surrogate endpoint trial data for decision making

One example: Is HbA1c a valid surrogate marker?

Limitations and threats for the use of surrogate endpoint data

## Rationale for the use of surrogate endpoints in phase III studies

- Efficacy data based on clinical endpoints require large sample size and long follow-up
  - Chronic diseases
  - Expected small effect sizes
  - Issue of costs
- As a consequence drugs are approved where efficacy is exclusively based on trials using surrogate endpoints

# Definition surrogate endpoint

*Clin Pharmacol Ther* 2001, 69(3):89-95.

- A surrogate endpoint is an indicator variable substituting for a clinically meaningful endpoint that reflects how a patient feels, functions or survives
- Surrogate endpoints include physiologic variables that are indicators of normal biological or pathogenic processes or pharmacological responses to a therapeutic intervention
- Biomarker

## Examples of surrogate endpoints

Surrogate endpoint

Clinical endpoint

---

Blood pressure

Stroke/ MI

Bone mineral density

Osteoporotic fracture

HIV-1 RNA viral load

AIDS/ AIDS related death

# Statistical surrogate marker validation

*Prentice Stat Med 1989;8:431–440*

- Treatment ( $Z$ ) has a significant impact on the surrogate endpoint ( $S_j$ ) and on the true endpoint ( $T_j$ )
- Surrogate endpoint has a significant impact on the true endpoint
- The full effect of treatment upon the true endpoint is captured by the surrogate  $Z \longrightarrow S_j \longrightarrow T_j$
- The more likely situation  $Z \longrightarrow S_j \longrightarrow T_j$



## An example: HbA1c in type II diabetes

- Should sitagliptin, a dipeptidyl peptidase-IV-inhibitor for the TX of type 2 diabetes in patients insufficiently controlled by metformin be used
- No evidence on the efficacy of sitagliptin to lower patient important endpoints or long-time safety

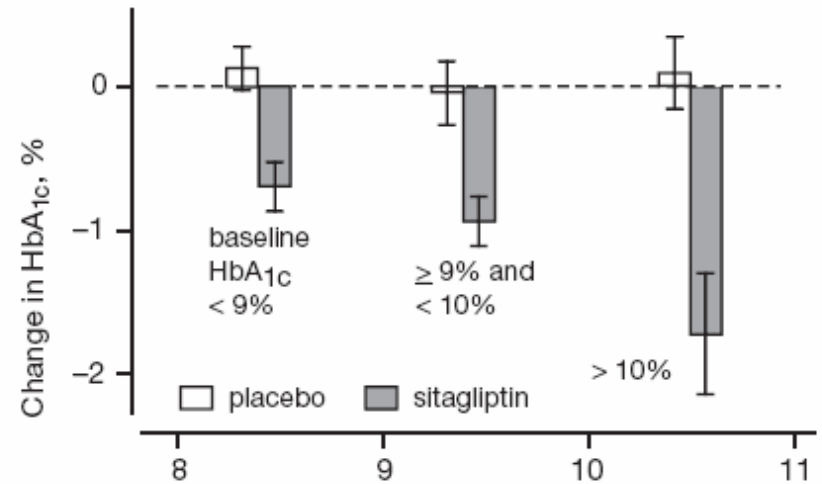
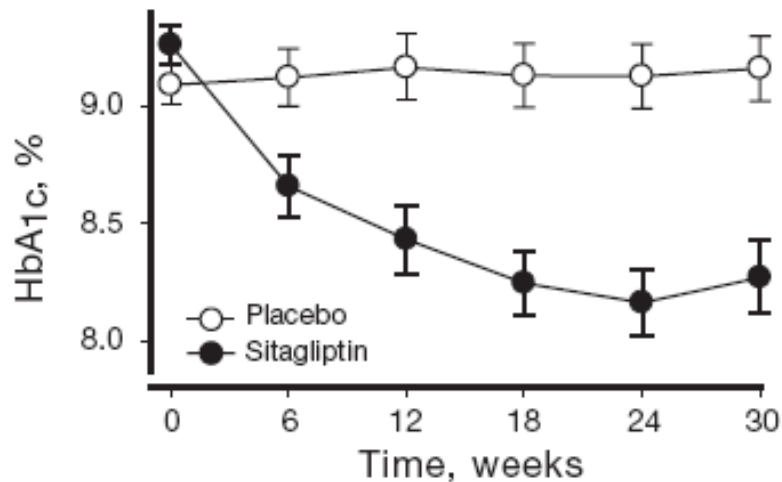
# A medical breakthrough to merit explicit reliance on surrogate outcome?





# Mean change in HbA1c between sitagliptin and placebo in patients treated with metformin

*Raz I et al Current Medical Research and Opinion 2008; 24: 537–50*



Number of patients in the trial: 190

## The clinical questions

- What are the criteria to judge the validity of evidence from surrogate endpoint studies?
- When should we base treatment decisions on the evidence from surrogate endpoint studies?
- Is HbA1c a useful surrogate marker to evaluate the efficacy of a glucose lowering agent in patients at very high risk for a MI or death from MI?

## 1. Validity criteria

Consistency of the association between the surrogate and clinical endpoint

Is there a strong, independent, consistent association between the surrogate endpoint and the clinical endpoint?

## 1.1 Strength of the association

- Yes, there is a strong association between HbA1c and risk of fatal MI

RR of overall death per 1 % increase in HbA1c:

Men 1.24 (CI, 1.14 - 1.34;  $P < 0.001$ )

Women 1.28 (CI, 1.06 - 1.32;  $P < 0.001$ )

*Norfolk Study: Ann Intern Med 2004;141:413*

## 1.2 Consistency of the association

Was there a consistent association between the surrogate endpoint and the clinical endpoint across several studies?

- Several studies have consistently shown an association between the level of HbA1c and subsequent risk of death or MI

*Kuusitu J et al Finish Cohort: Diabetes 1994: 43: 960*

## 2. Validity criteria Between drug class effect

Is there evidence from randomised controlled trials in other drug classes that improvement in the surrogate endpoint has consistently led to improvement in patient-important outcomes?

# Consistency of the association of HbA1c modification and myocardial infarction across drug classes *UKPDS Lancet 1998; 352: 837*

## Insuline /metformin vs metformin

- Surrogate outcome:
- HbA1c over 10 years 11% reduction
  - Intensive group 7.0% (6.2–8.2)
  - Conventional group 7.9% (6.9–8.8)

# Consistency of the association of HbA1c modification and myocardial infarction

*UKPDS Lancet 1998; 352: 837*

## Clinical outcome:

- Diabetes related endpoints
  - RRR 12% (95% CI 1–21,  $p=0.029$ )
- Any diabetes related death
  - RRR 10% (–11 to 27,  $p=0.34$ )
- All cause mortality
  - RRR 6% (–10 to 20,  $p=0.44$ )

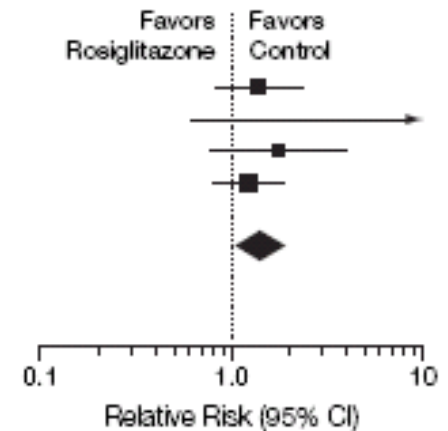


# Long-term Risk of Cardiovascular Events With Rosiglitazone A Meta-analysis *JAMA. 2007;298:1189-95*

## Myocardial Infarction

Source	No. of Events/Total (%)		Weight, %	Relative Risk (95% CI)
	Rosiglitazone	Control		
Kahn et al, <sup>7</sup> 2006	24/1456 (1.8)	34/2895 (1.2)	31.52	1.40 (0.84-2.36)
Dargie et al, <sup>18</sup> 2007	5/110 (4.5)	0/114 (0)	0.68	11.40 (0.64-203.69)
Gerstein et al, <sup>5</sup> 2006	16/2635 (0.6)	9/2634 (0.3)	12.47	1.78 (0.79-4.01)
Home et al, <sup>12</sup> 2007	49/2220 (2.2)	40/2227 (1.8)	55.33	1.23 (0.81-1.86)
<b>Total (95% CI)</b>	<b>6421</b>	<b>7870</b>	<b>100.00</b>	<b>1.42 (1.06-1.91)</b>

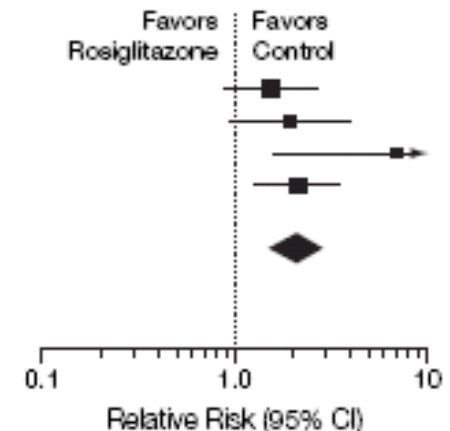
Total events: 94 (rosiglitazone), 83 (control)  
 Test for heterogeneity:  $\chi^2_3 = 2.77$  ( $P = .43$ ),  $I^2 = 0\%$   
 Tests for overall effect:  $Z = 2.33$  ( $P = .02$ )



## Heart Failure

Source	No. of Events/Total (%)		Weight, %	Relative Risk (95% CI)
	Rosiglitazone	Control		
Kahn et al, <sup>7</sup> 2006	22/1456 (1.5)	28/2895 (1.0)	35.68	1.56 (0.90-2.72)
Dargie et al, <sup>18</sup> 2007	19/110 (17)	10/114 (8.8)	18.70	1.97 (0.96-4.04)
Gerstein et al, <sup>5</sup> 2006	14/2635 (0.5)	2/2634 (0.1)	3.81	7.00 (1.59-30.76)
Home et al, <sup>12</sup> 2007	47/2220 (2.1)	22/2227 (1.0)	41.82	2.14 (1.30-3.54)
<b>Total (95% CI)</b>	<b>6421</b>	<b>7870</b>	<b>100.00</b>	<b>2.09 (1.52-2.88)</b>

Total events: 102 (rosiglitazone), 62 (control)  
 Test for heterogeneity:  $\chi^2_3 = 3.65$  ( $P = .30$ ),  $I^2 = 17.8\%$   
 Tests for overall effect:  $Z = 4.51$  ( $P = .00001$ )



### 3. Validity criteria

#### Within drug class effect

Is there evidence from randomised trials in the same drug class that improvement in the surrogate endpoint has consistently led to improvement in the target outcome?

No!

## 4. How to interpret results from surrogate studies?

How large, precise, and lasting was the treatment effect of the surrogate marker?

- Effect size
- Precision (95% confidence interval)
- Effect should be sufficiently lasting to be relevant for a clinical effect

## Mean HbA1c reduction & extrapolated benefit

- Mean HbA1c reduction of sitagliptin from a meta-analysis:  $-0.74\%$  (95%CI  $-0.84$  to  $-0.63$ ) *Amori JAMA 2007*
- May relate to a reduction of diabetes related endpoints of (*UKPDS Lancet 1998*):
  - RRR 12% (95% CI 1–21,  $p=0.029$ )
  - AR 40.9 vs 46.0 per 1000 patients years

## 5. Limitations and threats for the use of surrogate endpoint data

# Selected examples of applied validity criteria for the critical evaluation of studies using surrogate endpoints

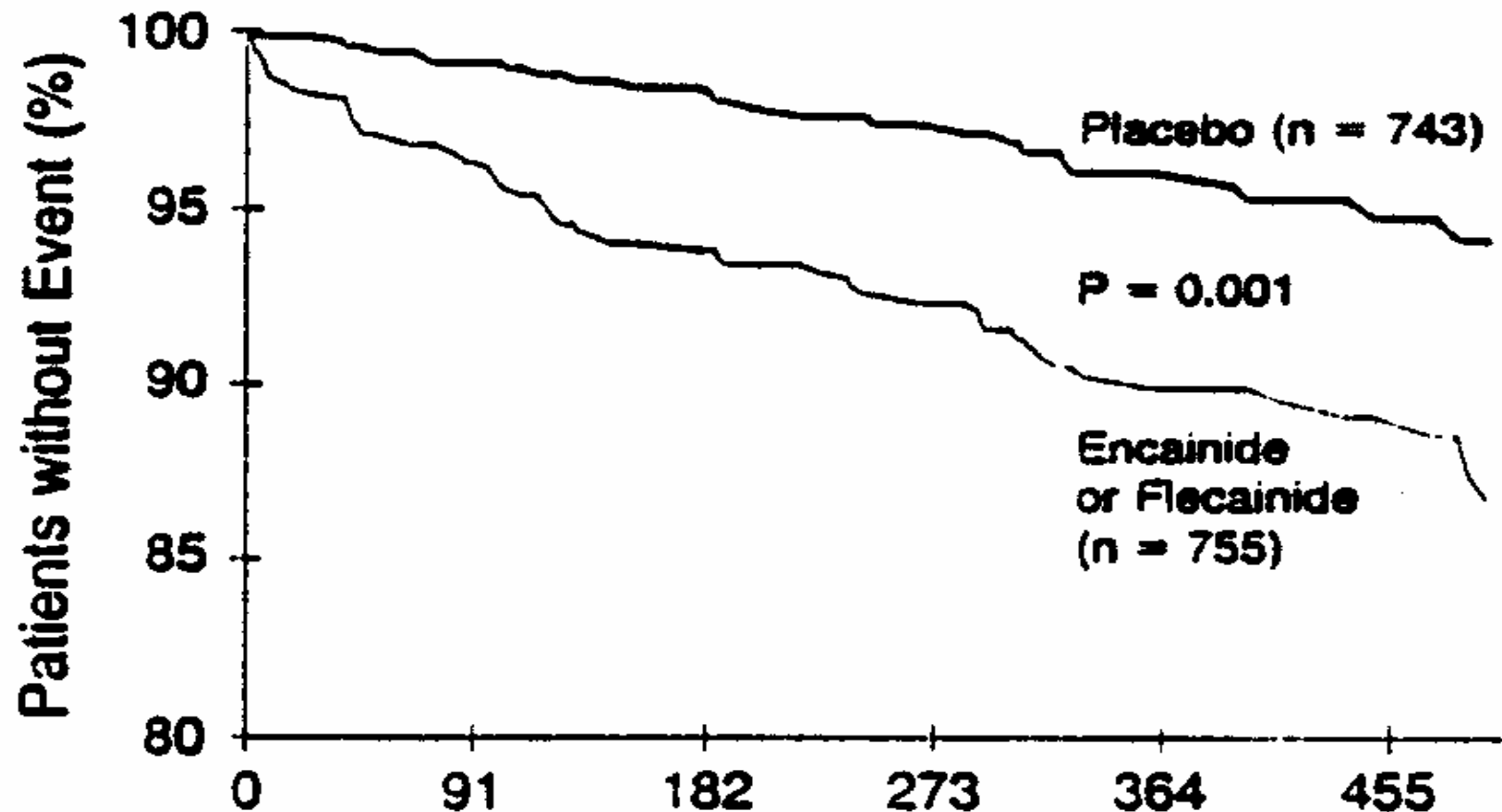
Substance	Surrogate marker	Clinical End point	Criteria 1 Association surrogate/ outcome	Criteria 2 Inter class effect	Criteria 3 Class effect
Sitagliptin	HbA1c	MI/death	Yes	No	No
Rosiglitazone	HbA1c	MI/death	Yes	No	No
Darunavir	CD4, viral load	AIDS or death	Yes	Yes	Yes
ACE inhibitor	Albuminuria in type II diabetes	ESRD	Yes	No	Yes
Rosuvastatin	Cholesterol	MI/death	Yes	No	Yes?

## Examples with biased conclusions from surrogate endpoint studies

Drug	Surrogate endpoint	Clinical endpoint
Encainamide, flecainamid	Complex VAS (Lown IV B)	Sudden death
Milrinone, epopostrol	Exercise capacity	Heart failure and death
Sodium fluoride	Bone mineral density	Fractures
Estrogen / progestin	Total/LDL cholesterol	Myocardial infarction



Probability of survival in post myocardial infarction patients treated with encainide, flecainide or placebo *N Engl J Med 1991;423:781*



## 6. When to rely on surrogate endpoints?

- Treatment decisions based on surrogate marker data are most likely to do more good than harm if:
  - All validity criteria apply
  - High baseline risk for the target disorder
  - Patient places high value on avoiding target disorder
  - No satisfactory alternative therapies exist

## Conclusions (HbA1c as surrogate marker) I

- HbA1c is surrogate marker not fulfilling all criteria and thus of limited use for TX decisions for drugs with no evidence from trials with clinical outcome data

## Conclusions (general) II

- In many fields of medicine reliance on surrogate endpoints is unavoidable (HIV infection!)
- For highly prevalent conditions reliance on surrogate endpoints has lead to harm of patients
- There is a need to keep the ballance between public, patients' and industry's interests to allow for early registration of truely innovative drugs based on surrogate outcome
- This should not preempt the medical community to insist on efficacy data from RCTs powered for clinical endpoints

## References

- HC Bucher et al. JAMA 1999; 282: 771
- Users' Guide to the medical Literature Manual for Evidence-Based Clinical Practice, 2nd edition  
Gordon Guyatt, Drummond Rennie Maureen O Meade, Deborah J Cook, New York 2008; pp 317-340. ISBN 978-0-07-159034-1
- HC Bucher Studien mit Surrogatendpunkten; Nutzen und Grenzen in der klinischen Entscheidungsfindung Internist (Berl) 2008; 49(6):681-687