Clinical endpoints

Erik Christensen

Associate Professor, Consultant Physician, MD, Dr. Med. Sci.

Department of Gastroenterology and Endocrinology

Bispebjerg Hospital

Copenhagen, Denmark

Clinical endpoint – an event used as marker of course of disease

- Like everything else, diseases develop in time.
- Description of the course in time is an important aspect of characterizing diseases, including the effect of therapies.
- A detailed description of the course of disease may be very complex
- Accordingly the problem is being dealt with in simpler terms: e.g. the **time from randomization** in an RCT **to an event or endpoint** of interest like death in survival analysis.

The outcome variable (endpoint) depends 1) on the disease and 2) the potential effect of the therapy

- Disease characteristics:
  - Steadily progressive (e.g. cancer)
    - Time to death, time to relapse, time to complication …
  - Acute and reversible (e.g. infection)
    - Days of incapacitation, duration of hospital stay …
  - Symptom in chronic disease in stable phase (e.g. pain in rheumatoid arthritis)
    - Pain score on visual analogue scale (VAS), duration of pain episodes …
  - Severity of chronic disease (e.g. diabetes mellitus)
    - HgbA1c level, occurrence of complications, renal insufficiency …

An example:

- Does eradication treatment work in duodenal ulcer?

Specify the four components (PICO) in the clinical question

- Patient or problem
- Intervention
- Comparison intervention (gold standard)
- Outcome or endpoint

Are there less recurrent ulcers in the first year after eradication therapy for duodenal ulcer disease than after ulcer healing with ranitidine for 6 weeks?

- **Patient or problem**
  - duodenal ulcer disease
- **Intervention**
  - eradication therapy - ranitidine, amoxycillin, metronidazole
- **Comparison intervention**
  - ranitidine
- **Outcome or end-point**
  - recurrent ulcer within the first year after treatment

Endpoints, which scale?

- Binary endpoint
  - alive/dead, recurrence yes/no, pain yes/no
  - If possible present failure time curves (’survival curves’)
    - Utilize both complete and censored observation times
- Ordinal scale
  - coma grade 1-4, Likert scale, CDAI (Crohns Disease Activity Index)
- Continuous scale
  - Blood pressure, HbA1c, serum cholesterol, microalbuminurea
- Global assessment
  - QALY, VAS, ADL

Likert scale (ordinal scale)

For each of the statements below, please indicate the extent of your agreement or disagreement by placing a tick in the appropriate column	5 Strongly agree	4 Agree	3 Neither agree nor disagree	2 Disagree	1 Strongly disagree
1. Criminals convicted of murder should be hanged					
2. Trial by jury should be abolished					

Crohn's Disease Activity Index (CDAI)

Clinical or laboratory variable	Weighting factor
Number of liquid or soft stools each day	x 2
Abdominal pain (graded from 0-3 on severity) each day	x 5
General well being: 0 (well) to 4 (terrible) each day	x 7
Presence of complications (fissures, fistulae, fever, arthralgia …)	x 20
Taking opiates for diarrhea	x 30
Presence of abdominal mass (0 none, 2 questionable, 5 definite)	x 10
Hematocrit of <0.47 in men and <0.42 in women	x 6
Percentage deviation from standard weight	x 1

Remission of Crohn's disease: CDAI < 150.
Severe disease CDAI > 450
Response of a therapy: a fall of the CDAI of more than 70 points

Quality-adjusted life year (QALY)

- A measure of disease burden, including both the quality and the quantity of life lived
- The QALY is based on the number of years of life that would be added by the intervention. Each year in perfect health is assigned the value of 1.0 down to a value of 0.0 for death.

Visual analogue scale (VAS)

No pain

Maximal pain

NB! – comparisons within the same individual only

Activities of Daily Living (ADL)

- A term used in healthcare to refer to daily self-care activities within an individual's place of residence, in outdoor environments, or both. Basic ADL (BADL) consist of self-care tasks, including:
  - Personal hygiene and grooming
  - Dressing and undressing
  - Self feeding
  - Functional transfers (Getting from bed to wheelchair etc.)
  - Bowel and bladder management
  - Ambulation

Endpoints, what do we want to assess?

- Hard endpoints
  - dead, social pension, fracture, helicobacter pylori present in gastric biopsy
- Soft endpoints
  - patients or doctors assessment i.e. pain, delusions, work capacity

Clinically relevant effect and surrogate endpoints?

- Clinically relevant effect
  - AMI, cerebrovascular insult, mortality, recurrent ulcer, pain or pain score, clinical score, quality of life score, etc.
  - Composite or combined endpoints
- Surrogate endpoint
  - blood glucose, blood pressure, peak-flow, serum cholesterol, microalbumin excretion in urine, CD4+ lymphocytes

Even hard endpoints should be carefully specified

- Death – a hard endpoint?
- Overall mortality?
- Disease related mortality?
- Procedure related mortality?
- Disease and procedure related mortality?

Disease and procedure related events

- Precise definitions in the protocol of events as endpoints
- Independent committee should evaluate possible events in relation to disease or procedure

Endpoint in phase 1-4 studies (example: cancer)

- Phase 1 studies
  - Surrogate endpoints (toxicity, side effects, tumour shrinkage)
- Phase 2 studies
  - Surrogate endpoints (dose, side effects, tumour shrinkage)
  - Clinically relevant endpoints – possibly (symptoms)
- Phase 3 studies
  - Surrogate end-points
  - Clinically relevant endpoints – preferred (remission)
- Phase 4 studies
  - Clinically relevant endpoints mainly (remission, survival)

Clinically relevant endpoints

All endpoints must be clearly defined in the protocol

- Primary endpoint
  - Only one primary endpoint
- Secondary endpoints
  - 2 or 3 are acceptable - priority given
- Composite endpoint
  - Combination of primary and secondary endpoints
- Possible surrogate endpoints
  - may be included for investigatory reasons
  - for comparison with primary or secondary endpoints

Important: Limit the number of endpoints

Surrogate endpoints - definitions

- A laboratory measurement or a physical sign in which changes induced by therapy are expected to reflect changes in a clinically meaningful endpoint (Temple 1995)
- An observed variable that relates in some way to the variable of primary interest (Hillis 1989)
- A response variable for which a test of the null-hypothesis - is also a valid test of the corresponding null-hypothesis based on true endpoints (Prentice 1989)

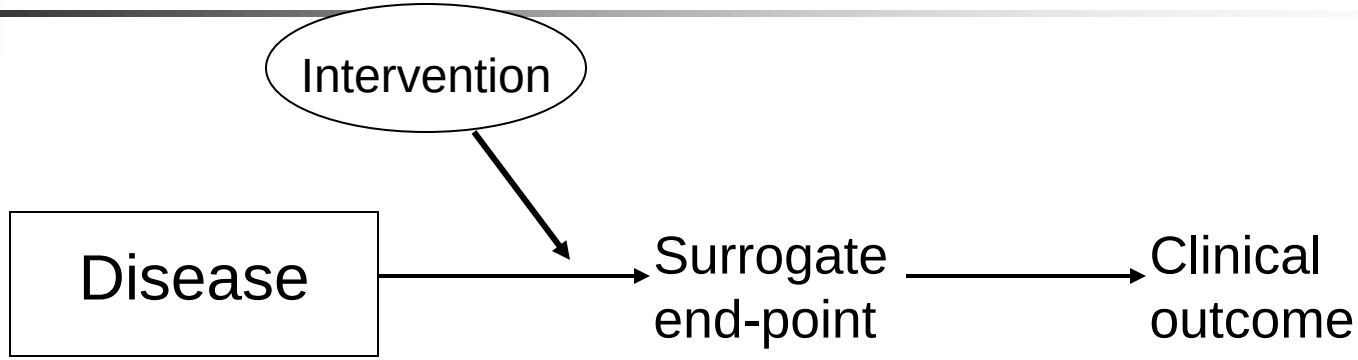
Why surrogate endpoints?

- Economic reasons
- Practical reasons
- Ethical reasons
- Scientific “precision”
- Statistical reasons

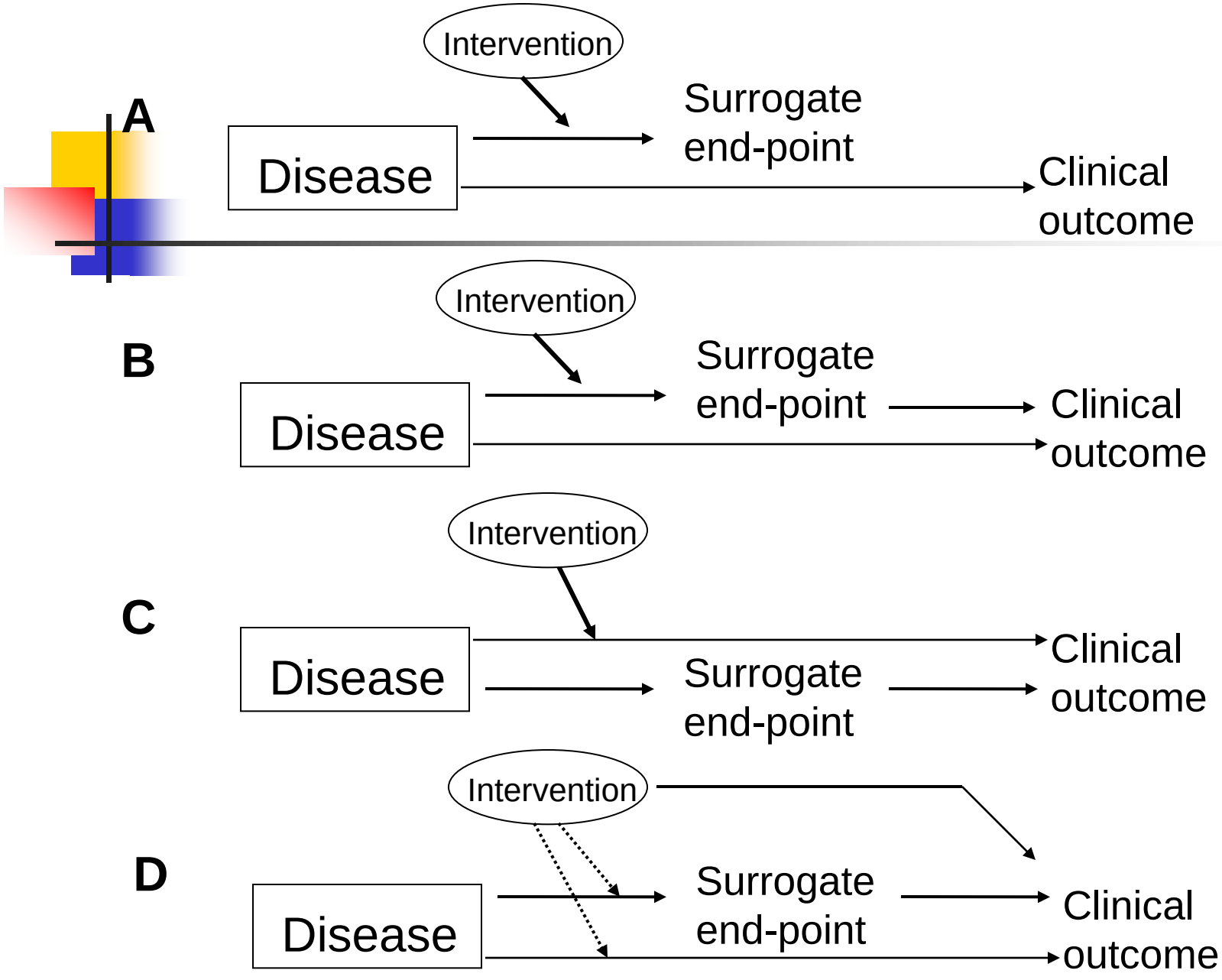
Requirements for surrogate endpoints

- Prognostic marker - analytical epidemiology
  - true marker or confounder?
- Biologic marker
  - etiologic role
  - pathophysiologic role
  - close causal relation to clinically relevant endpoint
- Statistical marker
  - more common than clinically relevant endpoint
  - correlates closely to clinically relevant endpoint

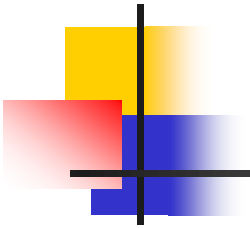
Optimal surrogate endpoint



- The effect of the intervention on the surrogate endpoint predicts the effect the clinical outcome
- The surrogate endpoint correlates with the true clinical outcome
- The surrogate endpoint fully captures the net effect on the clinical outcome


A
Intervention
Disease
Surrogate
end-point
Clinical
outcome
B
Intervention
Disease
Surrogate
end-point
Clinical
outcome
C
Intervention
Disease
Clinical
outcome
Surrogate
end-point
D
Intervention
Disease
Surrogate
end-point
Clinical
outcome

Selection of surrogate endpoints



- Prognostic factors from epidemiological studies
- Etiologic factors influenced by drug
- Pathophysiologic factors influenced by mechanism of drug action
- Prognostic factors from randomized clinical trials

Validation of surrogate end-points Boissel 1992

- Only epidemiological evidence available: *not sufficient*
- Only data on etiology and/or pathophysiology influenced by mechanism of action available: *not sufficient*
- Epidemiological data and pathophysiological data influenced by mechanism of action available: *possibly a surrogate and point*
- Epidemiological data and data from RCT with clinically relevant endpoints available: *possibly*
- Only data from RCT with clinically relevant endpoints available: *possibly*

The biomarker-surrogacy evaluation schema

Marissa N Lassere. Stat Methods Med Res. 2008;17:303-40.

- Study design criterion
  0 to 5 points
- Target outcome criterion
  0 to 5 points
- Statistical evaluation criterion
  0 to 5 points
- Penalties
  – 1 to – 3
- Level 1 (score 13-15) and Level 2 (score 10-12) are called ‘surrogates’, lower levels ‘biomarkers’

Composite endpoint

- Death or some other worst outcome combined with
- Other elements
  - that are clinically relevant
  - but some kind of surrogate for the worst outcome
  - i.e. pathophysiologically related to worst outcome

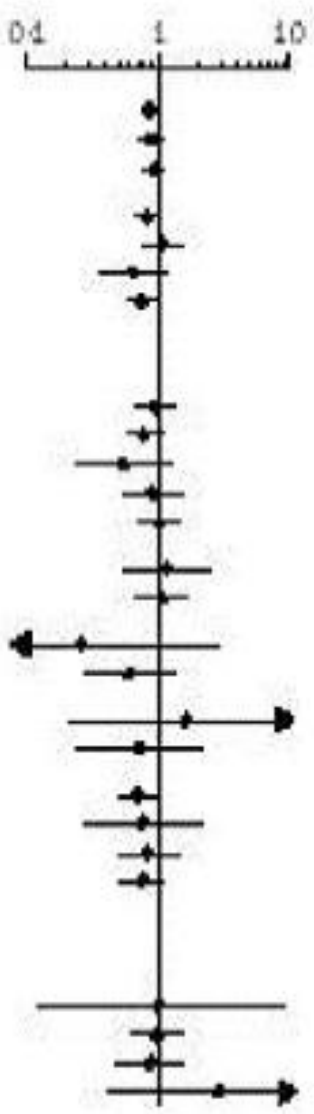
Composite endpoint

- Combination of
  - Primary endpoint
  - Secondary endpoint
  - Tertiary endpoint
  - etc.
  - etc.

- **UK Prospective Diabetes Study (**UKPDS**) Group.** Lancet. 1998;352:837-53.
  - Sudden death
  - Death from hyperglycemia
  - Death from hypoglycemia
  - Myocardial infarction
  - Angina
  - etc., etc.
  - Amputation (minimum one digit)
  - Retinopathy, photocoagulation
  - Blindness in one eye

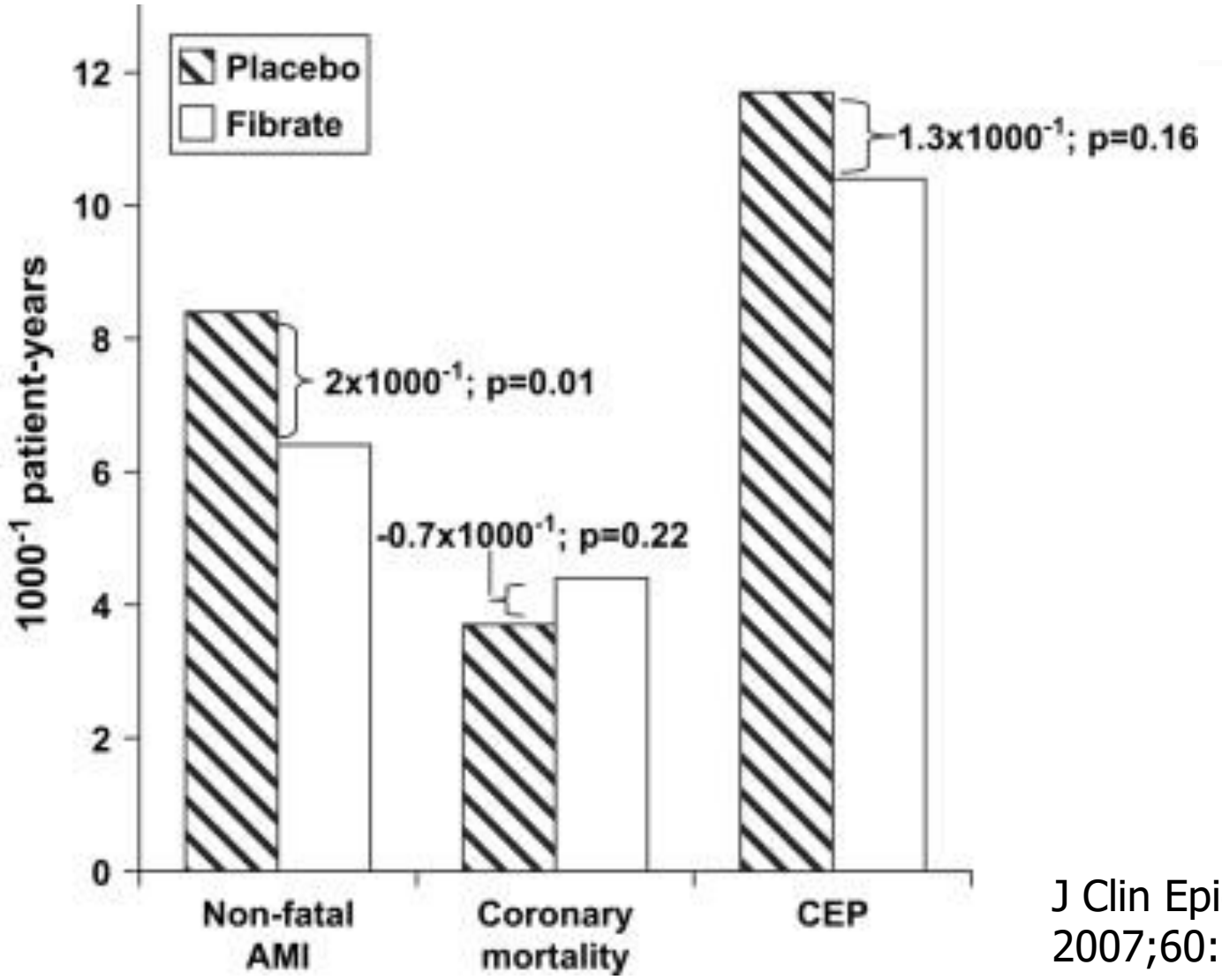
UK Prospective Diabetes Study (UKPDS**) Group.** Lancet. 1998;352:837-53.

	Patients with clinical endpoints		Absolute risk: events per 1000 patient-years		Log-rank p	RR for intensive policy (CI)
	Intensive (n=2729)	Conventional (n=1138)	Intensive	Conventional		
AGGREGATE ENDPOINT						
Any diabetes-related endpoint	563	438	40·9	46·0	0·029	0·88 (0·79–0·99)
Diabetes-related deaths	265	129	10·4	11·5	0·34	0·90 (0·73–1·11)
All-cause mortality	469	213	17·9	18·9	0·44	0·94 (0·80–1·10)
Myocardial infarction	387	186	14·7	17·4	0·052	0·84 (0·71–1·00)
Stroke	148	55	5·6	5·0	0·52	1·11 (0·81–1·51)
Amputation or death from PVD	29	18	1·1	1·6	0·15	0·65 (0·36–1·18)
Microvascular	225	121	8·6	11·4	0·0099	0·75 (0·60–0·93)
SINGLE ENDPOINTS						
Fatal myocardial infarction	207	90	7·6	8·0	0·63	0·94 (0·68–1·30)
Non-fatal myocardial infarction	197	101	7·5	9·5	0·057	0·79 (0·58–1·09)
Fatal: sudden death	24	18	0·9	1·6	0·047	0·54 (0·24–1·21)
Heart failure	80	36	3·0	3·3	0·69	0·91 (0·54–1·52)
Angina	177	72	6·8	6·7	0·91	1·02 (0·71–1·46)
Fatal stroke	43	15	1·6	1·3	0·60	1·17 (0·54–2·54)
Non-fatal stroke	114	44	4·3	4·0	0·72	1·07 (0·68–1·69)
Death from peripheral vascular disease	2	3	0·1	0·3	0·12	0·28 (0·03–2·77)
Amputation	27	18	1·0	1·6	0·099	0·61 (0·28–1·33)
Death from renal disease	8	2	0·3	0·2	0·53	1·63 (0·21–12·48)
Renal failure	16	9	0·6	0·8	0·45	0·73 (0·25–2·14)
Retinal photocoagulation	207	117	7·9	11·0	0·0031	0·71 (0·53–0·96)
Vitreous haemorrhage	19	10	0·7	0·9	0·51	0·77 (0·28–2·11)
Blind in one eye	78	38	2·9	3·5	0·39	0·84 (0·51–1·40)
Cataract extraction	149	80	5·6	7·4	0·046	0·76 (0·53–1·08)
Death from hyperglycaemia	0	1	0	0·1		
Death from hypoglycaemia	1	0	0	0		
Fatal accident	5	2	0·2	0·2	0·99	1·01 (0·12–8·70)
Death from cancer	120	50	4·4	4·4	0·92	0·98 (0·64–1·52)
Death from any other specific cause	65	30	2·4	2·7	0·57	0·88 (0·50–1·56)
Death from unknown cause	14	2	0·5	0·2	0·14	2·68 (0·41–20·19)



RR=relative risk. 95% CI for aggregate and 99% CI for single endpoints. PVD=peripheral vascular disease.

Composite endpoint (CEP)



J Clin Epidemiol. 2007;60:651-7

Composite endpoints – requirements

- All components are prespecified and clinically relevant
- All components must represent aspects of the same pathophysiological process
- Relative risk reduction for endpoints of same magnitude
- Effect of treatment about the same for alle components
- Should mirror the clinical spectrum of outcomes

These requirements are seldom fulfilled

Composite Endpoints (CEP) - weaknesses

- The treatment effect may be difficult to interpret, because the various components in CEP are not equivalent
- If the less serious endpoints dominate in the CEP, a treatment effect in CEP may be seen, even if the more serious endpoints in CEP are nearly equally distributed

Disease severity index – a possible alternative to composite endpoints

- Use a disease severity index (e.g. CDAI, or a prognostic index PI demonstrated to correlate with a hard clinically relevant endpoint (e.g. death)).
- A PI is the weighted sum of the patients prognostic variables at the time in question
- Measure the index in all patients at various time intervals after randomization
- Compare the PI curves for the treatments statistically.
- Advantage: greater statistical power because all patients contribute. Analysis does not depend on endpoints, which may be scarce.

Conclusion

- Relevant clinical endpoints – preferred
- What is clinically relevant
  - Some soft endpoints may be more relevant than hard endpoints
- Surrogate endpoints
  - Necessary for developmental studies
  - But for clinical use – be very cautious
- Precise definitions of endpoints are mandatory
- If relevant clinical endpoints are scarce - consider using a disease severity index