The STROBE statement Amendment proposals

Bendix Carstensen Steno Diabetes Center A/S
Gentofte, Denmark
www.biostat.ku.dk/~bxc/

DCRC, Banff,
March 2011

STROBE

STrengthening the **R**eporting of
OBservational studies in **E**pidemiology [1, 2]

STROBE

STrengthening the **R**eporting of
OBservational studies in **E**pidemiology [1, 2]

- Contains guidelines on **reporting** studies.

STROBE

STrengthening the **R**eporting of
OBservational studies in **E**pidemiology [1, 2]

- Contains guidelines on **reporting** studies.
- But little in the way of
  data collection and **analysis**.

STROBE

STrengthening the **R**eporting of
OBservational studies in **E**pidemiology [1, 2]

- Contains guidelines on **reporting** studies.
- But little in the way of
  data collection and **analysis**.
  — by the very name of the statement. . .

STROBE

STrengthening the **R**eporting of
OBservational studies in **E**pidemiology [1, 2]

- Contains guidelines on **reporting** studies.
- But little in the way of
  data collection and **analysis**.
  — by the very name of the statement. . .
- Particular issues arise in analysing data from follow-up based on registers:

STROBE

STrengthening the **R**eporting of
OBservational studies in **E**pidemiology [1, 2]

- Contains guidelines on **reporting** studies.
- But little in the way of
  data collection and **analysis**.
  — by the very name of the statement. . .
- Particular issues arise in analysing data from follow-up based on registers:
  - Person follow-up in continuous time
    (all dates known)
  - Person exposure in continuous time.

STROBE

STrengthening the **R**eporting of
OBservational studies in **E**pidemiology [1, 2]

- Contains guidelines on **reporting** studies.
- But little in the way of
  data collection and **analysis**.
  — by the very name of the statement. . .
- Particular issues arise in analysing data from follow-up based on registers:
  - Person follow-up in continuous time
    (all dates known)
  - Person exposure in continuous time.
- . . . as opposed to detailed clinical data collected at visits to clinics/doctors.

Follow-up studies

- STROBE has very little about long-term follow-up studies and the necessary reporting of this.

Follow-up studies

- STROBE has very little about long-term follow-up studies and the necessary reporting of this.
- Probabilists consider the disease-process as dependent on a **filtration**, $\mathcal{F}_t$:

$$\mathrm{P}\{(d, y)(t)|\mathcal{F}_t\}$$

Follow-up studies

- STROBE has very little about long-term follow-up studies and the necessary reporting of this.
- Probabilists consider the disease-process as dependent on a **filtration**, $\mathcal{F}_t$:

$$\mathrm{P}\{(d, y)(t)|\mathcal{F}_t\}$$

— which just means that what goes on at time t only depends on the past.

Follow-up studies

- STROBE has very little about long-term follow-up studies and the necessary reporting of this.
- Probabilists consider the disease-process as dependent on a **filtration**, $\mathcal{F}_t$:

$$\mathrm{P}\{(d, y)(t)|\mathcal{F}_t\}$$

  — which just means that what goes on at time t only depends on the past.
- Implies that we should look at follow-up in small pieces
  — in the limit at every single time of follow up.

Follow-up studies

- Every assumption about the disease process at a time t should be independent of whether a person at observation is lost to follow-up tomorrow or not.

Follow-up studies

- Every assumption about the disease process at a time t should be independent of whether a person at observation is lost to follow-up tomorrow or not.
- Implications for:

Follow-up studies

- Every assumption about the disease process at a time t should be independent of whether a person at observation is lost to follow-up tomorrow or not.
- Implications for:
  - Data selection
    (matching, choice of "comparators", selection of follow-up)

Follow-up studies

- Every assumption about the disease process at a time t should be independent of whether a person at observation is lost to follow-up tomorrow or not.
- Implications for:
  - Data selection
    (matching, choice of "comparators", selection of follow-up)
  - Covariate definition

Follow-up studies

- Every assumption about the disease process at a time t should be independent of whether a person at observation is lost to follow-up tomorrow or not.
- Implications for:
  - Data selection
    (matching, choice of "comparators", selection of follow-up)
  - Covariate definition
  - Model construction

Follow-up studies

- Every assumption about the disease process at a time t should be independent of whether a person at observation is lost to follow-up tomorrow or not.
- Implications for:
  - Data selection
    (matching, choice of "comparators", selection of follow-up)
  - Covariate definition
  - Model construction
  - Parameter interpretation

Data selection

- Persons:

Data selection

- Persons:
  - Preferably the entire population.
    Anything less gives less precision (power)
    . . . but the important figure is the no. exposed cases.

Data selection

- Persons:
  - Preferably the entire population.
    Anything less gives less precision (power)
    . . . but the important figure is the no. exposed cases.
  - Select only on variables known at (intended) entry to the follow-up.

Data selection

- Persons:
  - Preferably the entire population.
    Anything less gives less precision (power)
    . . . but the important figure is the no. exposed cases.
  - Select only on variables known at (intended) entry to the follow-up.
- Time window in study:

Data selection

- Persons:
  - Preferably the entire population.
    Anything less gives less precision (power)
    . . . but the important figure is the no. exposed cases.
  - Select only on variables known at (intended) entry to the follow-up.
- Time window in study:
  - As large as possible

Data selection

- Persons:
  - Preferably the entire population.
    Anything less gives less precision (power)
    . . . but the important figure is the no. exposed cases.
  - Select only on variables known at (intended) entry to the follow-up.
- Time window in study:
  - As large as possible
  - Selected independent of future exposure status

Data selection

- Persons:
  - Preferably the entire population.
    Anything less gives less precision (power)
    . . . but the important figure is the no. exposed cases.
  - Select only on variables known at (intended) entry to the follow-up.
- Time window in study:
  - As large as possible
  - Selected independent of future exposure status
- **STROBE:** Already a note on informative loss to follow-up.
  "State clearly if persons are excluded from follow-up prior to death or end of follow-up."

Covariate definition

- Only what can be known at the particular time of follow-up.

Covariate definition

- Only what can be known at the particular time of follow-up.
- But not restricted to the **status** at that time.

Covariate definition

- Only what can be known at the particular time of follow-up.
- But not restricted to the **status** at that time.
- The entire **history** of the person should be included:

Covariate definition

- Only what can be known at the particular time of follow-up.
- But not restricted to the **status** at that time.
- The entire **history** of the person should be included:
  - Cumulative drug exposure (length of window)

Covariate definition

- Only what can be known at the particular time of follow-up.
- But not restricted to the **status** at that time.
- The entire **history** of the person should be included:
  - Cumulative drug exposure (length of window)
  - Cumulative time on drug

Covariate definition

- Only what can be known at the particular time of follow-up.
- But not restricted to the **status** at that time.
- The entire **history** of the person should be included:
  - Cumulative drug exposure (length of window)
  - Cumulative time on drug
  - Time since start of drug

Covariate definition

- Only what can be known at the particular time of follow-up.
- But not restricted to the **status** at that time.
- The entire **history** of the person should be included:
  - Cumulative drug exposure (length of window)
  - Cumulative time on drug
  - Time since start of drug
  - Current drug exposure

Covariate definition

- Only what can be known at the particular time of follow-up.
- But not restricted to the **status** at that time.
- The entire **history** of the person should be included:
  - Cumulative drug exposure (length of window)
  - Cumulative time on drug
  - Time since start of drug
  - Current drug exposure
- Unrelated to data and data availability

Covariate definition

- Only what can be known at the particular time of follow-up.
- But not restricted to the **status** at that time.
- The entire **history** of the person should be included:
  - Cumulative drug exposure (length of window)
  - Cumulative time on drug
  - Time since start of drug
  - Current drug exposure
- Unrelated to data and data availability
- An ideal specification of relevant determinants

Data definition

— essentially practical covariate **encoding**.

- Different from definition of covariates:

Data definition

— essentially practical covariate **encoding**.

- Different from definition of covariates:
- Includes decisions on how to make approximations based on available data:

Data definition

— essentially practical covariate **encoding**.

- Different from definition of covariates:
- Includes decisions on how to make approximations based on available data:
  - time

Data definition

— essentially practical covariate **encoding**.

- Different from definition of covariates:
- Includes decisions on how to make approximations based on available data:
  - time
  - dose

Data definition

— essentially practical covariate **encoding**.

- Different from definition of covariates:
- Includes decisions on how to make approximations based on available data:
  - time
  - dose
- Limitations from data

Data definition

— essentially practical covariate **encoding**.

- Different from definition of covariates:
- Includes decisions on how to make approximations based on available data:
  - time
  - dose
- Limitations from data
- Limitations from analytical approach

Data definition

— essentially practical covariate **encoding**.

- Different from definition of covariates:
- Includes decisions on how to make approximations based on available data:
  - time
  - dose
- Limitations from data
- Limitations from analytical approach
- Single day as analytical unit is hardly feasible.

Data definition

— essentially practical covariate **encoding**.

- Different from definition of covariates:
- Includes decisions on how to make approximations based on available data:
  - time
  - dose
- Limitations from data
- Limitations from analytical approach
- Single day as analytical unit is hardly feasible.
- **STROBE:** "State the aims of covariate definition (exposure) independent of the data available, preferably based on a biological / clinical hypothesis"

Model specification

- Definition of covariates in a model induces a way to look at the determinants of the disease process:

Model specification

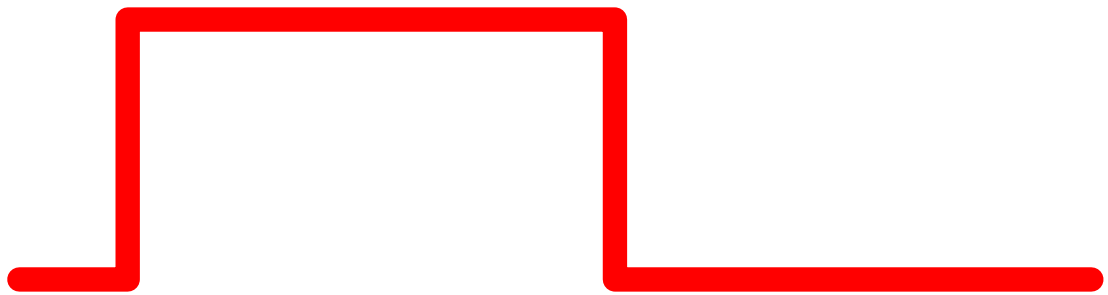
- Definition of covariates in a model induces a way to look at the determinants of the disease process:
- Include only indicator of current use of drug A in the model:

Model specification

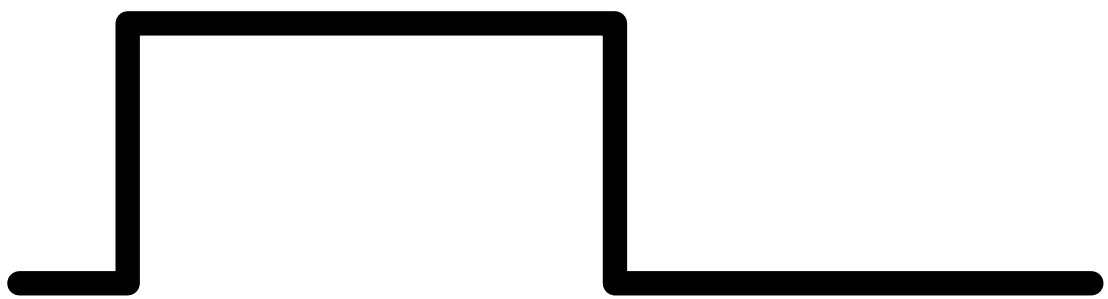
- Definition of covariates in a model induces a way to look at the determinants of the disease process:
- Include only indicator of current use of drug A in the model:
  - Risk of outcome increases immediately at drug start

Model specification

- Definition of covariates in a model induces a way to look at the determinants of the disease process:
- Include only indicator of current use of drug A in the model:
  - Risk of outcome increases immediately at drug start
  - Risk decreases to precisely the previous level at drug cessation

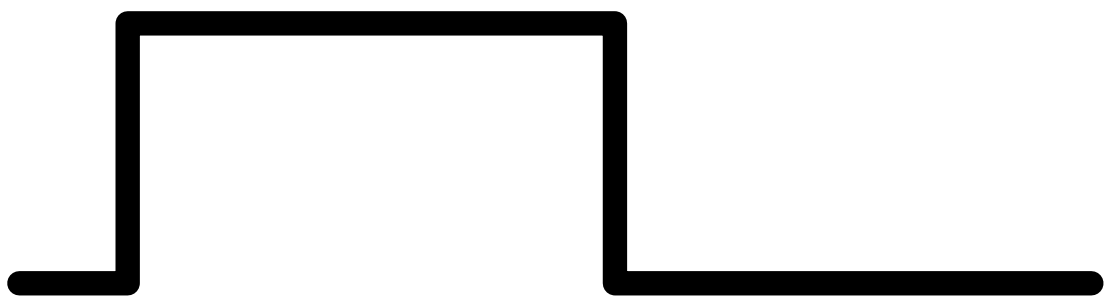


"Drug effect"



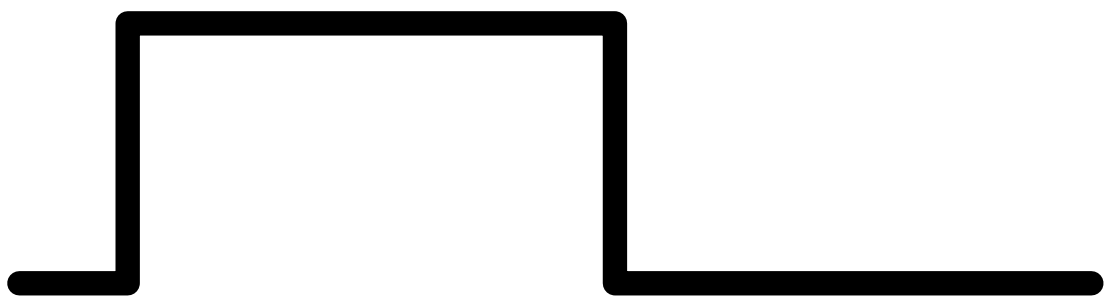
Drug exposure

"Drug effect"



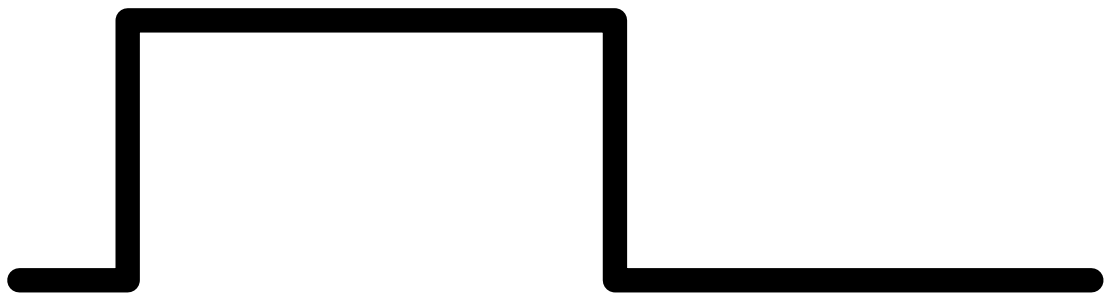
Drug exposure

"Drug effect"



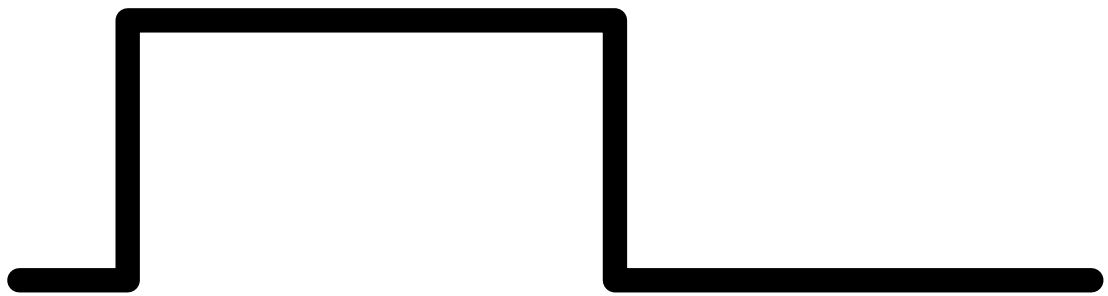
Drug exposure

"Drug effect"

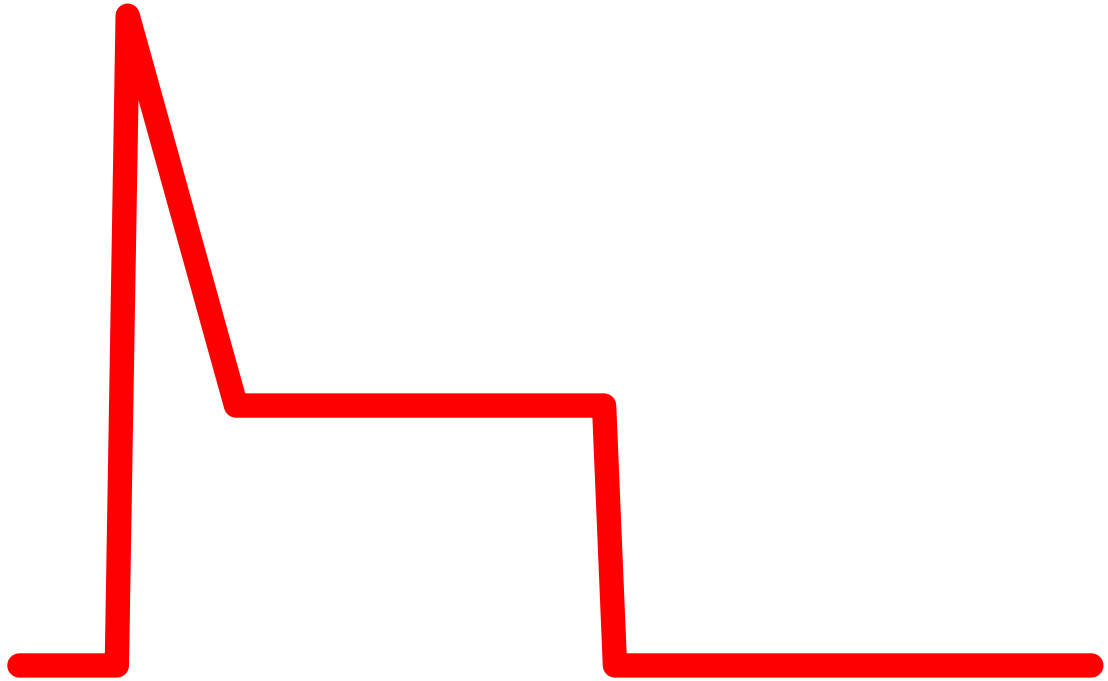


Drug exposure

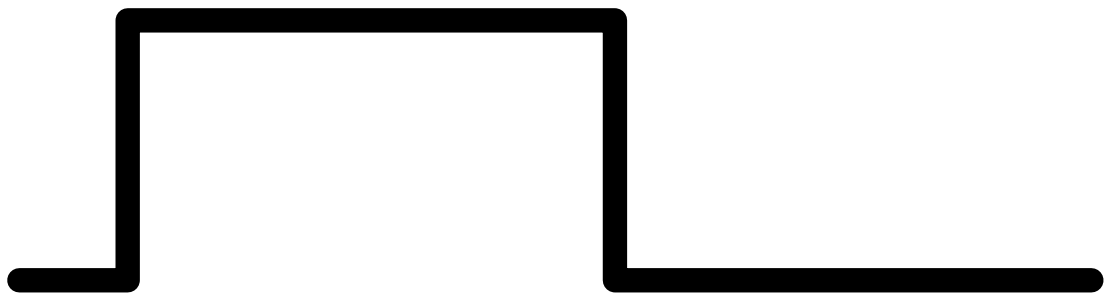
Indication effect



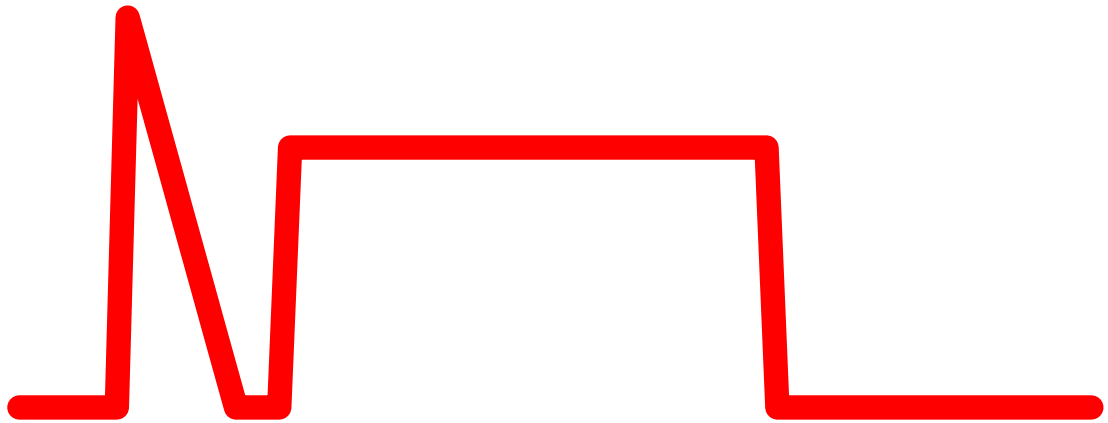
Drug exposure



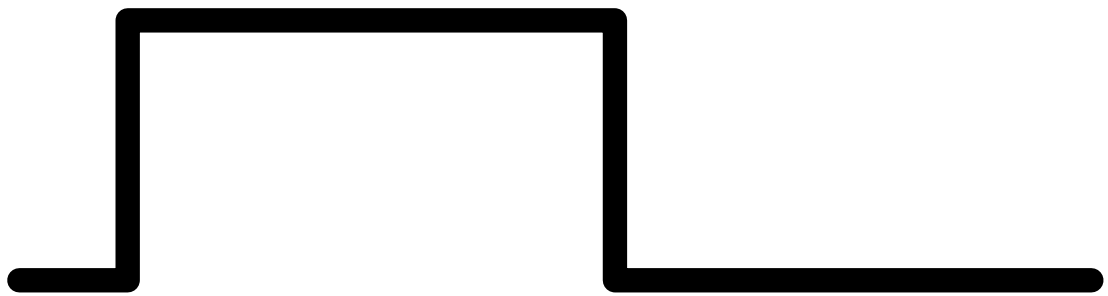
"Drug effect"



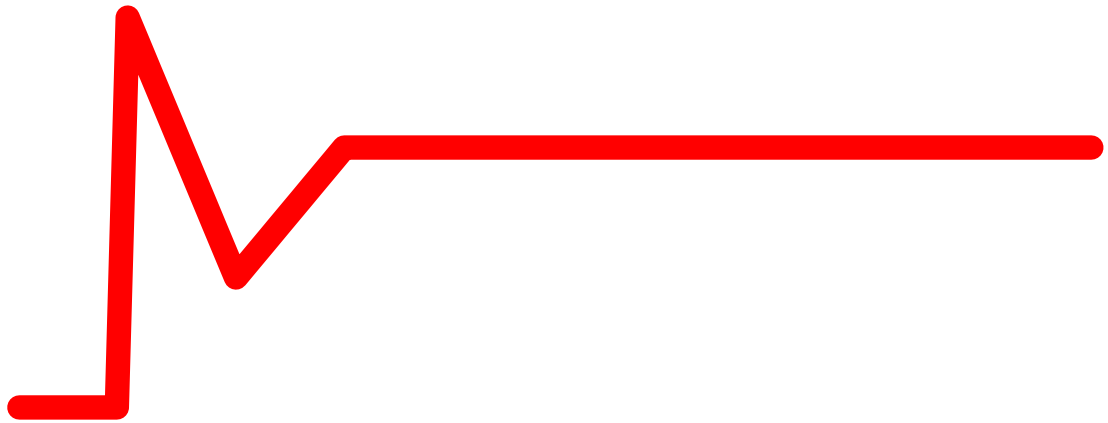
Drug exposure



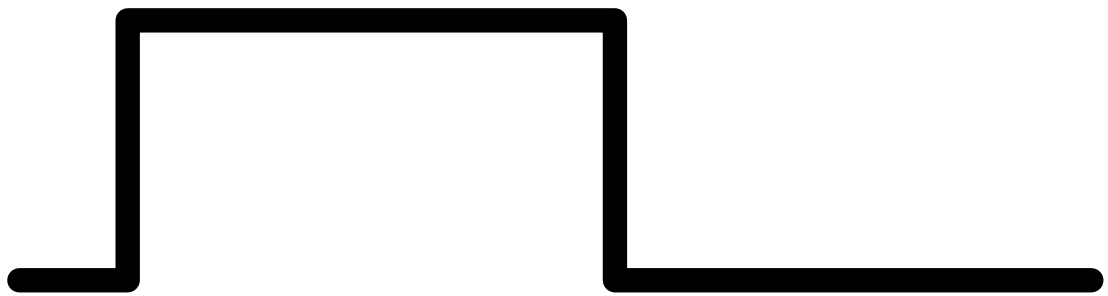
"Drug effect"



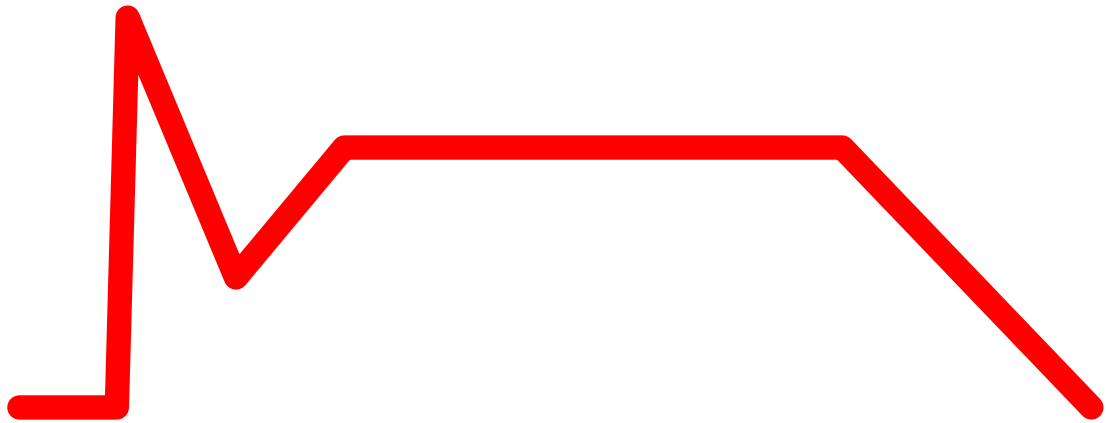
Drug exposure



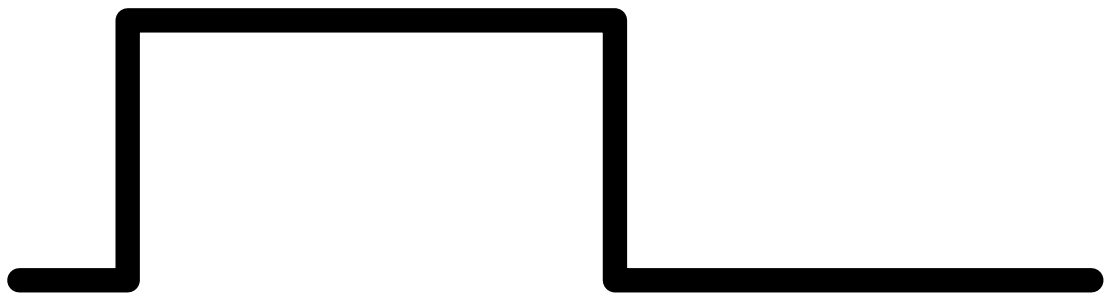
"Drug effect"



Drug exposure



"Drug effect"



Drug exposure

Model specification

- Is this biologically/clinically plausible?

Model specification

- Is this biologically/clinically plausible?
- An approximation to something biologically/clinically plausible?

Model specification

- Is this biologically/clinically plausible?
- An approximation to something biologically/clinically plausible?
- No better approximation available?

Model specification

- Is this biologically/clinically plausible?
- An approximation to something biologically/clinically plausible?
- No better approximation available?
- **STROBE:** "Explain precisely how the anticipated temporal relationship between drug exposure and effect is modeled."

Model specification

- Is this biologically/clinically plausible?
- An approximation to something biologically/clinically plausible?
- No better approximation available?
- **STROBE:** "Explain precisely how the anticipated temporal relationship between drug exposure and effect is modeled."
- We assumed that initiation of XX caused an immediate increase in rates of cancer which remained constant

.

Model specification

- Is this biologically/clinically plausible?
- An approximation to something biologically/clinically plausible?
- No better approximation available?
- **STROBE:** "Explain precisely how the anticipated temporal relationship between drug exposure and effect is modeled."
- We assumed that initiation of XX caused an immediate increase in rates of cancer which remained constant, and only were present for the duration of drug exposure.

Model specification

- Definition of covariates in a model induces a way to look at the determinants of the disease process:

Model specification

- Definition of covariates in a model induces a way to look at the determinants of the disease process:
- A single covariate (duration *e.g.*) may be modeled using more parameters.

Model specification

- Definition of covariates in a model induces a way to look at the determinants of the disease process:
- A single covariate (duration *e.g.*) may be modeled using more parameters.
- Drug exposure in a data base is not only (an approximate) measurement of drug exposure, but also of processes around it — indication effects.

Parameter interpretation

- A set of covariates induces a model for the risk (rates) of the event (e.g. cancer death) as a function of these:

Parameter interpretation

- A set of covariates induces a model for the risk (rates) of the event (e.g. cancer death) as a function of these:
- If **rates** are shown, we must specify all covariates.

Parameter interpretation

- A set of covariates induces a model for the risk (rates) of the event (e.g. cancer death) as a function of these:
- If **rates** are shown, we must specify all covariates.
- Sometimes average rates for a certain joint distribution of covariates is given:

Parameter interpretation

- A set of covariates induces a model for the risk (rates) of the event (e.g. cancer death) as a function of these:
- If **rates** are shown, we must specify all covariates.
- Sometimes average rates for a certain joint distribution of covariates is given:
- Rates controlled for...

.

Parameter interpretation

- A set of covariates induces a model for the risk (rates) of the event (e.g. cancer death) as a function of these:
- If **rates** are shown, we must specify all covariates.
- Sometimes average rates for a certain joint distribution of covariates is given:
- Rates controlled for... is most often really Rates in a population with a covariate distribution as our study population....

Parameter interpretation

- When the intercept is in the reporting, specify **all** covariates.

Parameter interpretation

- When the intercept is in the reporting, specify **all** covariates.
- **STROBE:** "If absolute quantities (rates, risk, ...) are given, state the covariate values (or distribution of such) for which these are computed."

Parameter interpretation

- A set of covariates induces a model for the risk (rates) of the event (e.g. cancer death) as a function of these:

Parameter interpretation

- A set of covariates induces a model for the risk (rates) of the event (e.g. cancer death) as a function of these:
- If only **contrasts** (differences or ratios) are shown:

Parameter interpretation

- A set of covariates induces a model for the risk (rates) of the event (e.g. cancer death) as a function of these:
- If only **contrasts** (differences or ratios) are shown:
- The effect of a covariate assuming that all other are fixed.

So what is relevant for STROBE?

- Data base (why are data actually recorded in the first place)
- Data accuracy (exact date of purchase, date of cancer diagnosis only by month,. . .)
- Data rounding (drug exposure only computed every 3 or 6 months)
- Covariate definition at each point of follow-up
- Model assumptions about effects

Proposals for STROBE?

- "State clearly if persons are excluded from follow-up prior to death or end of follow-up."
- "State the aims of covariate definition (exposure) independent of the data available, preferably based on a biological / clinical hypothesis"
- "Explain precisely how the anticipated temporal relationship between drug exposure and effect is modeled."
- "If absolute quantities (rates, risk, ...) are given, state the covariate values (or distribution of such) for which these are computed."

References

J. P. Vandenbroucke, E. von Elm, D. G. Altman, P. C. Goetzsche, C. D. Mulrow, S. J. Pocock, C. Poole, J. J. Schlesselman, and M. Egger.
Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration.
PLoS Med., 4:e297, Oct 2007.

E. von Elm, D. G. Altman, M. Egger, S. J. Pocock, P. C. Goetzsche, and J. P. Vandenbroucke.
The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies.
PLoS Med., 4:e296, Oct 2007.