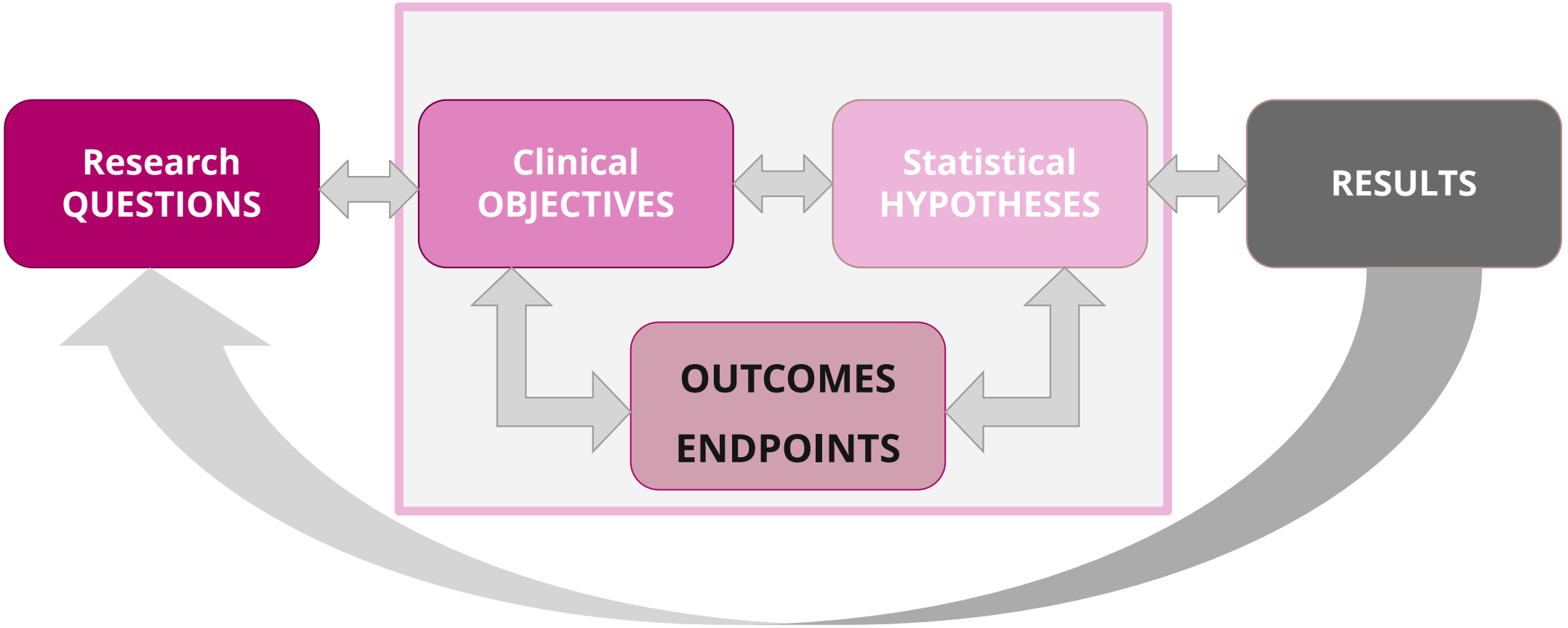
Endpoints in clinical trials and time-to-event data

Iveta Selingerova, Masaryk Memorial Cancer Institute

The main goals

1. Introduction to the terminology
2. Key definitions and endpoints
3. Endpoints in cancer trials
4. Binary outcome
5. Time-to-event data and survival analysis
6. Predictors (prognostic or predictive factors)
7. Interpretation of clinical trials results

From questions to results



Objectives/Aims

- Express the research question and summarize the goal of a clinical trial
- Should be concrete, be clearly specified
- Involve intent to do something with data obtained from clinical trial
- Expressed as a statement of purpose (e.g., to assess, to determine, to compare, to evaluate)
- Primary objectives
  - Answer the most important questions being asked by the trial
  - Drive statistical planning
- Secondary objectives
  - Answer other relevant questions about the trial

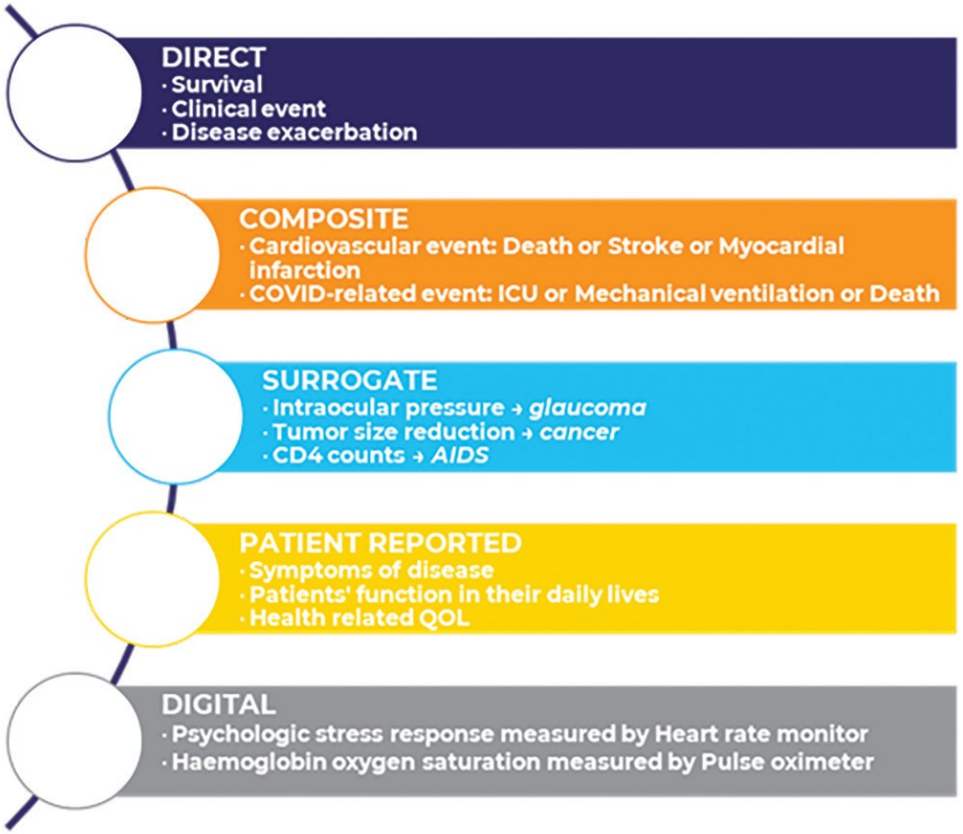
Endpoints/Outcomes

- The quantitative measurement or data point needed to meet the trial objectives
- Primary, secondary and exploratory endpoints
- The selection of endpoints is a complex process requiring a combination of clinical relevance and statistical reasoning
- Should be clinically relevant, precise, accurate, reliable, analyzable and interpretable
- Objective
  - can be measured without conditioning or influences
  - e.g., death, tumor response
- Subjective
  - most affected by individual interpretation
  - e.g., pain, depression or sleep quality

Clinical Distinctions Among Endpoints

- Direct: represent or characterize by themselves the clinical outcome of interest and directly measure the patient's clinical condition, function or survival.
- Surrogate: indicators closely correlated with clinically meaningful endpoints that can act as substitutes for those. They could be predictive of a clinical benefit in a shorter observation period.
- Composite: measure a unique effect by combining multiple endpoints.
- Patient reported: provided directly from the patient, without interpretation of the response by a doctor or anyone else.
- Digital: generated by a sensor, typically outside of a clinical setting during daily activities.

Clinical Distinctions Among Endpoints



Statistical Distinctions Among Endpoints

- Qualitative: data are characterized by impossibility of expressing them numerically
  - **Binary**: take one of two values
  - **Nominal**: can take on multiple values

  } individual categories cannot be compared or ordered
  - **Ordinal**: can take on multiple categories that can be sorted, however, it is not possible to evaluate distance between the categories
- Quantitative: data can be calculated or measured and expressed as a numerical value
  - **Discrete**: represented by a countable number of points
  - **Continuous**: can take on arbitrary values with a finite or infinite interval
- Time-to-event: data consist of two closely linked components - the time variable (quantitative) and the event indicator (binary)

Example of Outcome Types

- Qualitative:
  - **Binary**: Complete remission (yes vs no), Patient status (death vs alive)
  - **Nominal**: Treatment type, Adverse event type
  - **Ordinal**: Disease state (improved vs stable vs worsened), Neutropenia grade (1-4)
- Quantitative:
  - **Discrete**: pain scale (0=no pain, 10=worst pain imaginable)
  - **Continuous**: Gene expression, Tumor size in mm, Visual analog scale in mm, Quality of Life score
- Time-to-event: Time to death, Time to progression, Time to response

Common Endpoints in Cancer Clinical Trials

Serial No	Endpoint	Definition	Unique feature
1.	Overall survival	Time from randomization to death.	The 'gold standard' primary clinical endpoint.
2.	Progression-free survival	Time from randomization to disease progression or death, whichever comes first.	Used to assess therapies targeting advanced or metastatic malignancies.
3.	Time to progression	Time from randomization to disease progression.	Only uses time to progression and does not include time to death.
4.	Event-Free Survival	Time from randomization to disease progression, discontinuation of treatment for any reason, or death.	Used to evaluate highly toxic treatments.
5.	Disease-free survival	Time from randomization to disease recurrence.	Used to assess adjunctive and curative therapies.
6.	Time to Treatment Failure	Time from initiation of chemotherapy to premature discontinuation of treatment.	Used with other endpoints to assess reasons for discontinuing treatment.
7.	Time to Next Treatment	Time from initiation of treatment to beginning the next line of therapy.	Used as a meaningful endpoint for patients with low grade, incurable malignancies.
8.	Duration of Clinical Benefit	Time from randomization to progression or death in patients who had a complete or partial response or a stable disease for over 24 weeks.	Used in settings where disease stabilization is meaningful.
9.	Duration of Response	Time from randomization to progression or death in patients who had a complete or partial response.	Used to assess therapies for durable response.
10.	Objective Response Rate	Proportion of patients with partial or complete response to therapy.	Used to assess neoadjuvant therapies.
11.	Complete Response	Lack of detectable evidence of tumor.	Included as a major goal of multiple myeloma treatment.
12.	Pathological Complete Response	Lack of residual invasive cancer in resected breast tissue or regional lymph nodes.	Used in accelerated approval for neoadjuvant therapies targeting breast cancer.
13.	Disease Control Rate	Percentage of patients with complete response, partial response, or stable disease as a result of their therapy.	Used to assess the tumorstatic efficacy of a therapy.
14.	Clinical Benefit Rate	Percentage of patients with complete response, partial response, or at least months of stable disease as a result of their therapy.	Used to capture tumorstatic efficacy of a therapy and stable disease.
15.	Health-Related Quality of Life	Assessment of patient quality of life with respect to health status.	Used to directly measure patient quality of life.
16.	Milestone survival	Survival probability at a prespecified time point.	Used to evaluate a cross-section of OS data.

Delgado A, Guddati AK. Clinical endpoints in oncology - a primer. Am J Cancer Res. 2021

Common Endpoints in Cancer Clinical Trials

TABLE 2.5

Commonly used efficacy endpoints in oncology clinical trials: advantages and limitations

Endpoints	Definition	Advantages	Limitations
Overall survival (OS)	Time from randomization[a] until death from any cause	• Universally accepted measure of direct benefit • Easily and precisely measured	• May require a larger trial population and longer follow-up to show statistical difference between groups • May be affected by crossover or subsequent therapies • Includes deaths unrelated to cancer
Progression-free survival (PFS)	Time from randomization[a] until disease progression or death	• Requires small sample size and shorter follow-up time compared with OS • Includes measurement of stable disease (SD)	• Validation as a surrogate for survival can be difficult in some treatment settings • Not precisely measured (i.e., measurement may be subject to bias)
Time to progression (TTP)	Time from randomization[a] until objective tumor progression; does not include deaths	• Not affected by crossover or subsequent therapies • Generally based on objective and quantitative assessment	• Definition may vary among trials • Requires frequent radiologic or other assessments • Requires balanced timing of assessment among treatment arms
Time to treatment failure (TTF)	Time from randomization[a] to discontinuation of treatment for any reason, including disease progression, treatment toxicity, and death	• Useful in settings in which toxicity is potentially as serious as disease progression (e.g., allogeneic stem cell transplant)	• Does not adequately distinguish efficacy from other variables, such as toxicity
Event-free survival (EFS)	Time from randomization[a] to disease progression, death, or discontinuation of treatment for any reason (e.g., toxicity, patient preference, or initiation of a new treatment without documented progression)	• Similar to PFS; may be useful in evaluation of highly toxic therapies	• Initiation of next therapy is subjective. Generally not encouraged by regulatory agencies because it combines efficacy, toxicity, and patient withdrawal
Time to next treatment (TTNT)	Time from end of primary treatment to institution of next therapy	• For incurable diseases, may provide an endpoint meaningful to patients	• Not commonly used as a primary endpoint • Subject to variability in practice patterns
Objective response rate (ORR)	Proportion of patients with reduction in tumor burden of a predefined amount	• Can be assessed in single-arm trials • Requires a smaller population and can be assessed earlier, compared with survival trials	• Not a comprehensive measure of drug activity
Duration of response (DoR)	Time from documentation of tumor response to disease progression	• Effect is attributable directly to the drug, not the natural history of the disease	

Source: Adapted from U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER). Guidance for industry: clinical trial endpoints for the approval of cancer drugs and biologics. http://www.fda.gov/downloads/Drugs/…/Guidances/ucm071590.pdf. Published May 2007. Accessed February 22, 2016.

[a] Not all trials are randomized. In non-randomized trials, time from study enrollment is commonly used.

Halabi, Susan, and Stefan Michiels, *Textbook of clinical trials in oncology: a statistical perspective*.

Primary Endpoints

- The most important outcome
- Used to assess the primary objective of a trial
- Used to assess the effect of the study intervention
- Provide the basis for concluding whether the study met its objective
- Directly affect the design and sample size
- Most clinical studies have one primary outcome measure, but a clinical study may have more than one
- More primary endpoint: overall false-positive rate (α) should be allocated across endpoints (the split need not be even)

Binary Outcome

- take one of two values, usually either the presence or absence of a condition or the occurrence of an adverse event such as death or disease recurrence
- **Statistical measure:** probability estimated as proportion
- **Example of objective:** To compare presence of disease between treatment arms

	Presence	Absence	Total
Arm A	n_1	$N_A - n_1$	N_A
Arm B	n_2	$N_B - n_2$	N_B
Total	n	N-n	N

Binary Outcome

- Proportions: $p_1 = \frac{n_1}{N_A}$ $p_2 = \frac{n_2}{N_B}$

	Presence	Absence	Total
Arm A	p_1	$1-p_1$	1
Arm B	p_2	$1-p_2$	1
Total	p	1-p	1

Summary Statistics for Binary Outcomes

Name	Notation	Definition	Range	Equality point
Risk Difference	RD	$p_1 - p_2$	[−1,1]	0
Risk ratio/ Relative risk	RR	$\frac{p_1}{p_2}$	[0,∞)	1
Odds Ratio	OR	$\frac{\frac{p_1}{1-p_1}}{\frac{p_2}{1-p_2}}$	[0,∞)	1

Binary Outcome and other important terms

	Presence (P)	**Absence (A)**
Positivity (Pos)	True positive (TP)	False positive (FP)
Negativity (Neg)	False negative (FN)	True negative (TN)

Binary Outcome and other important terms

	Presence (P)	Absence (A)
Positivity (Pos)	True positive (TP)	False positive (FP)
Negativity (Neg)	False negative (FN)	True negative (TN)

- **Sensitivity** or True Positive Rate:

$$\frac{TP}{P} = \frac{TP}{TP + FN}$$

  - proportion of positive test in the group of people with presence of disease

Binary Outcome and other important terms

	Presence (P)	Absence (A)
Positivity (Pos)	True positive (TP)	False positive (FP)
Negativity (Neg)	False negative (FN)	True negative (TN)

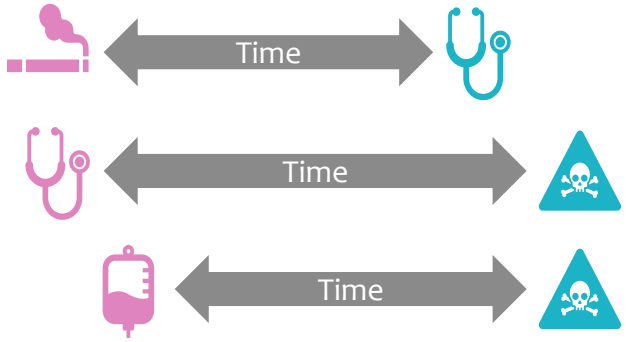
- **Specificity** or True Negative Rate:

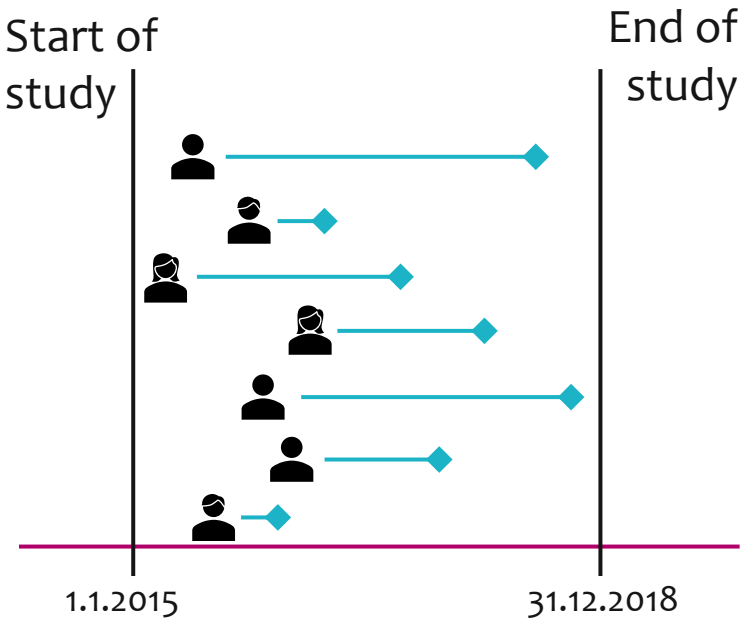
$$\frac{TN}{A} = \frac{TN}{FP + TN}$$

  - proportion of negative test in the group of people with absence of disease

Time-to-event data

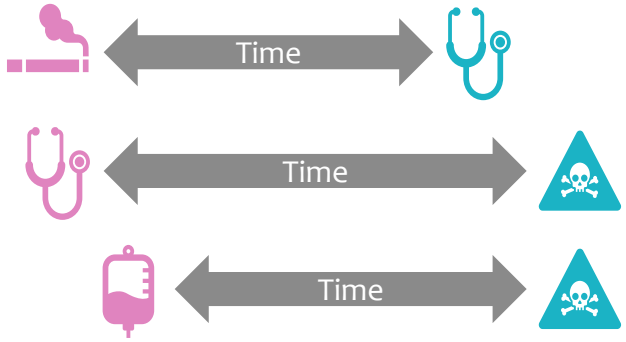
Describe time from well-defined time-origin until occurrence of event of interest

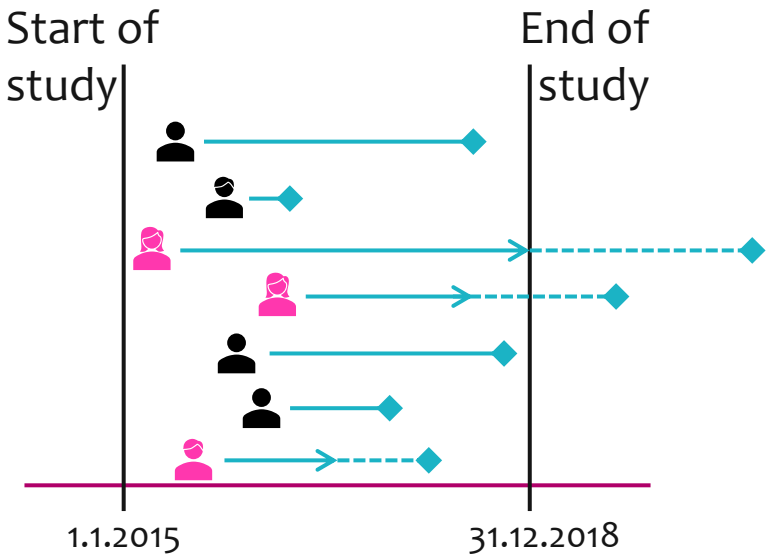




Time-to-event data

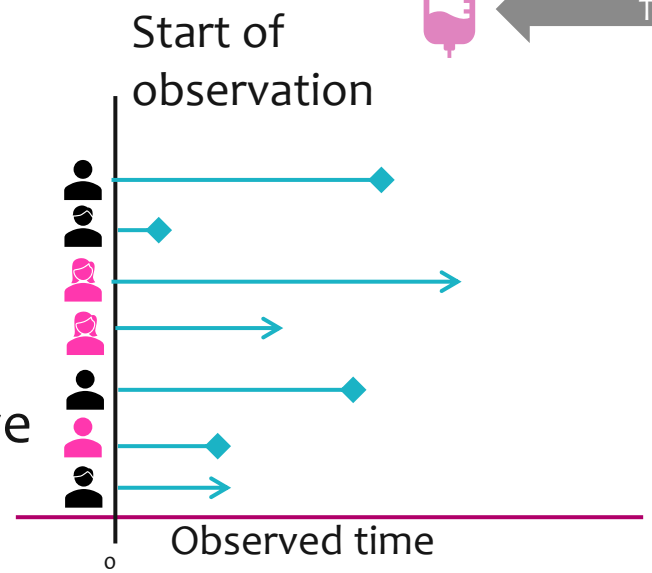
Describe time from well-defined time-origin until occurrence of event of interest





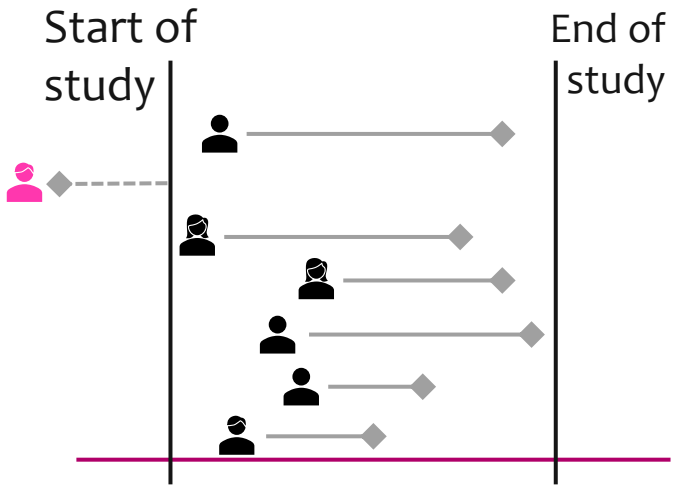
Right censoring

Assumption: Non-informative censoring

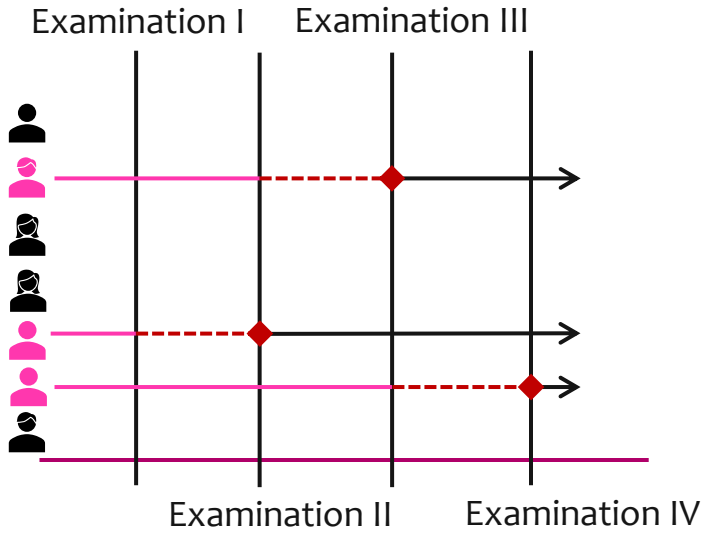


Censoring

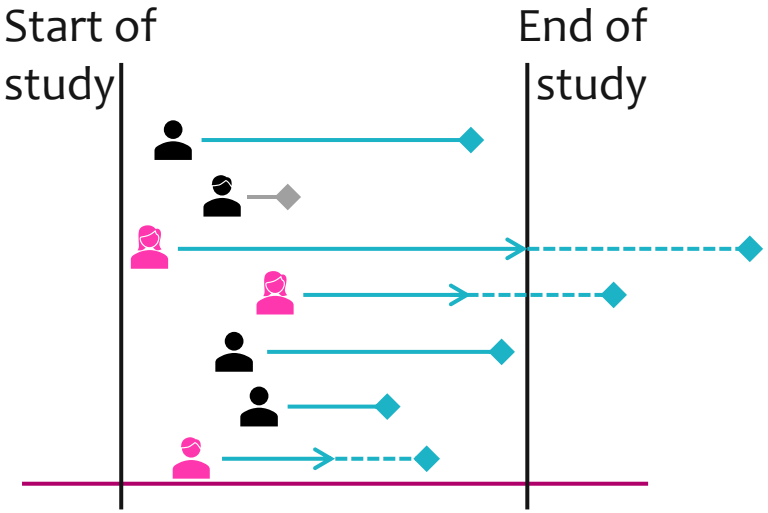
Left censoring



Interval censoring



Right censoring



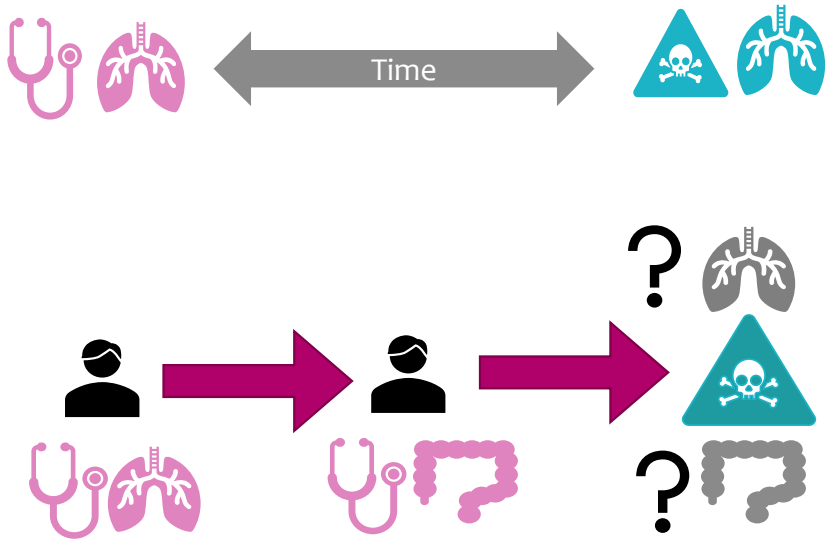
Censoring

- Censoring rate is a measure of study quality

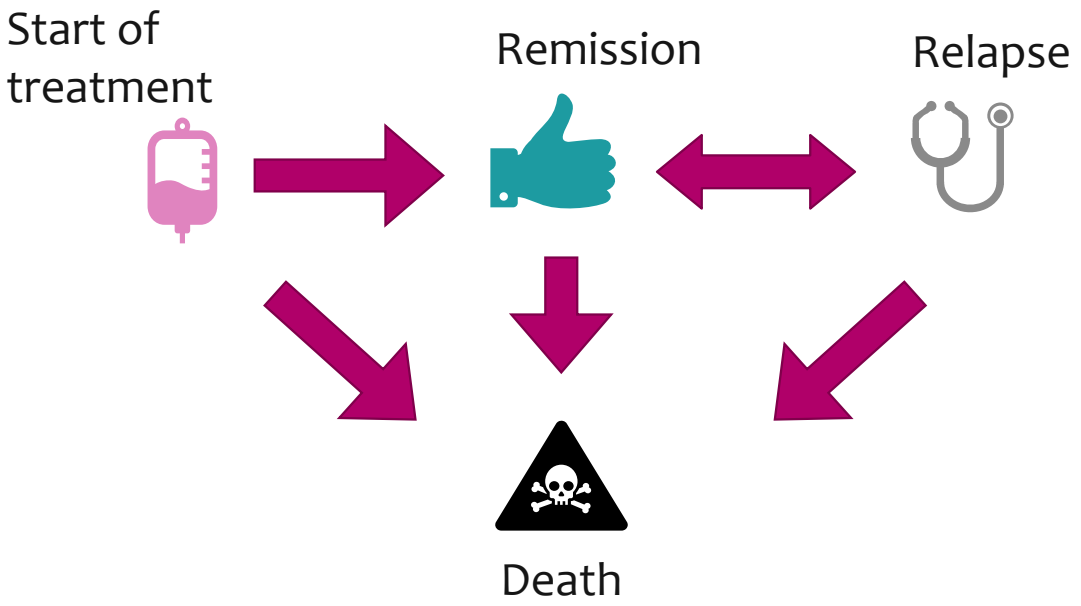
- **Other terminology:**
  - Truncation –data are observed only for individuals who have experienced a certain event, i.e., event time lies within a certain observational window are observed
  - Lead time – the time between the early diagnosis of a specific disease through screening and the time in which the diagnosis would have been established without screening (false impression of prolonged survival or slower progression of disease)

Specific Case of Time-to-event Data

Competing events



Recurrent events



Time-to-event endpoints

Table 2. DATECAN guidelines for clinical events to be included in the definitions of time-to-event end points in randomized clinical trials assessing treatments for breast cancer

Setting	Recommended Time-to-event end point	Causes of death included in definition					Clinical events included in definitions							
		From breast cancer	From non-breast cancer cause	Related to protocol treatment	From any cause	From unknown cause	Invasive ipsilateral breast tumor recurrence/ progression	Local invasive recurrence/ progression	Regional invasive recurrence/ progression (M+: regional progression)	Invasive contra lateral breast cancer	Appearance/ occurrence of metastases/ distant recurrence	Second primary invasive cancer (non-breast cancer)	Ipsilateral DCIS	Contra lateral DCIS
Non- metastatic	BCSS	X		NC										
	iDFS	X	X	X	X	X	X	X	X	X	X	X	X	X
	D-DFS	X	X	X	X	X					X			
	D-RFS	X	X	X	X	X					X			
	RFS	X	X	X	X	X	X	X	X		X		X	
	L-RFS	X	X	X	X	X	X	X	X				X	
	RFi	X					X	X	X		X		X	
	BCFi	X					X	X	X	X	X		X	X
	D-RFi	X									X			
Metastatic	PFS	X	X	X	X	X	NA	NA	X		X			
	TTP	X					NA	NA	X		X			

It was recommended not to include the following events in any of the time-to-event end points: loss to follow-up.
BCSS, breast cancer-specific survival; iDFS, invasive disease-free survival; D-DFS, distant disease-free survival; D-RFS, distant relapse-free survival; RFS, relapse-free survival; L-RFS, locoregional relapse-free survival; RFi, recurrence-free interval; BCFi, breast cancer-free interval; D-RFi, distant recurrence-free interval; PFS, progression-free survival; TTP, time-to-progression; NC, no consensus.

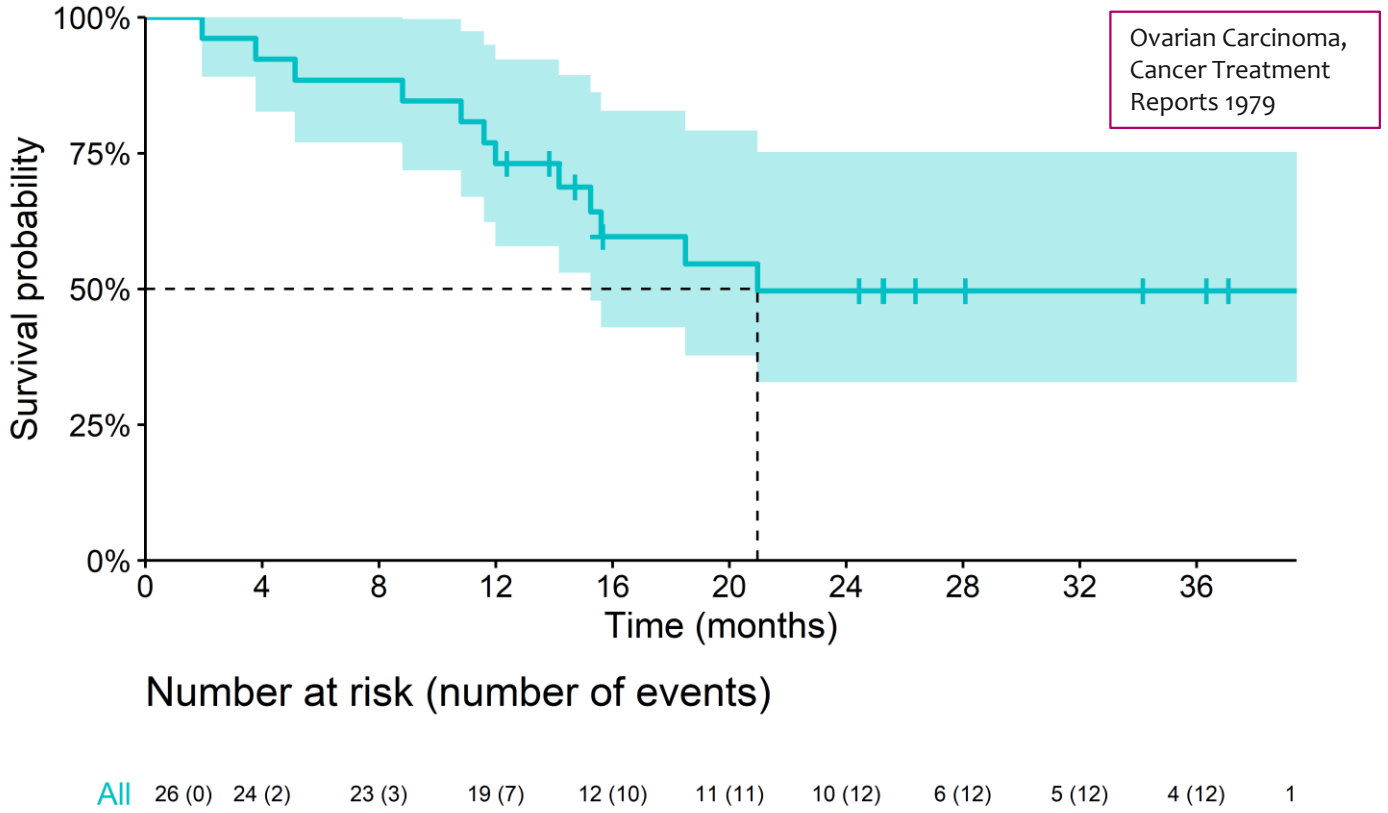
Guidelines for time-to-event end point definitions in breast cancer trials Annals of Oncology 2015

Basic Quantities of Survival Time

Survival function

The probability of an individual surviving beyond time point

Kaplan-Meier estimate of survival function



Follow-up

Table 1. Follow-up estimates stratified by the method used

Disease stage	Number of patients	Percentage of patients censored	Median follow-up in months					Median overall survival in months
			T-OBS	*T-CENS*	*T-END*	*KFT*	*Reverse-KM*	
I	62,427	85%	79	81	96	94	93	Not reached
II	11,402	63%	45	59	84	78	76	114
III	8,946	43%	41	79	156	144	129	65
IV	3,332	16%	7	55	132	132	152	8
Total	86,107	75%	66	78	96	96	92	320

- Description of how long the study was able to observe patients.
- It assesses the quality of the study, i.e., whether it was possible to capture a sufficient number of events (together with the proportion of censored data)

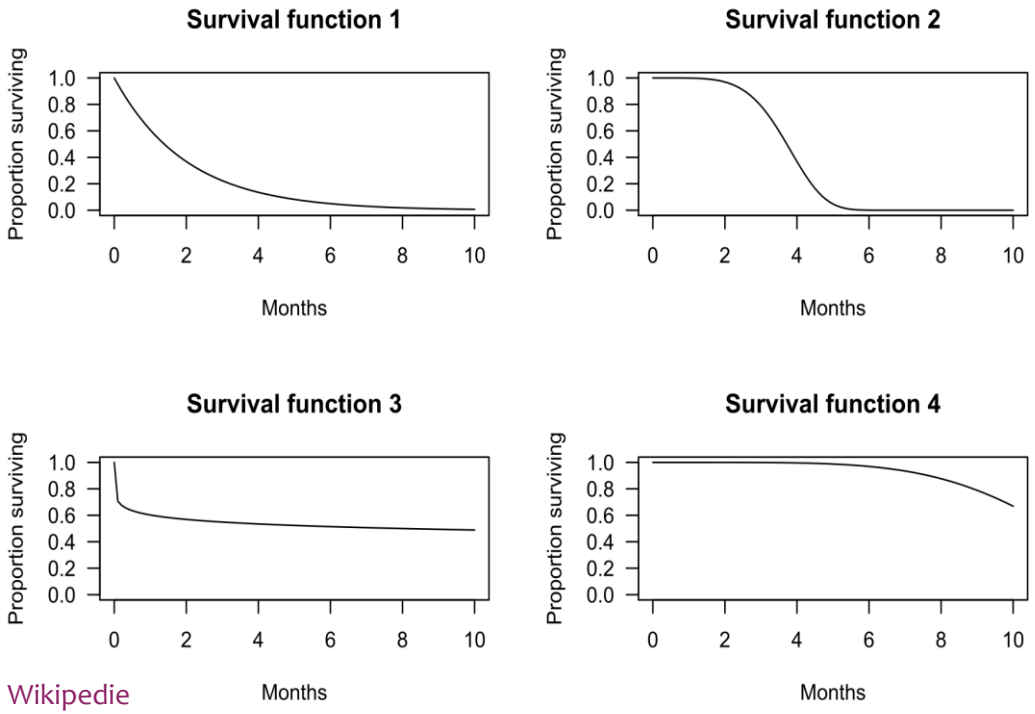
Methods to estimate follow-up

- **Observation time** (T-OBS) = T_2 - T_1
- **Time to censoring** (T-CENS) = T_2 - T_1 for surviving patients only
- **Time to end-of-study** (T-END) = T_3 - T_1
- **Known function time** (KFT) = T_2 - T_1 for surviving patients and T_3 - T_1 for deceased patients
- **Reverse Kaplan-Meier** (reverse-KM)

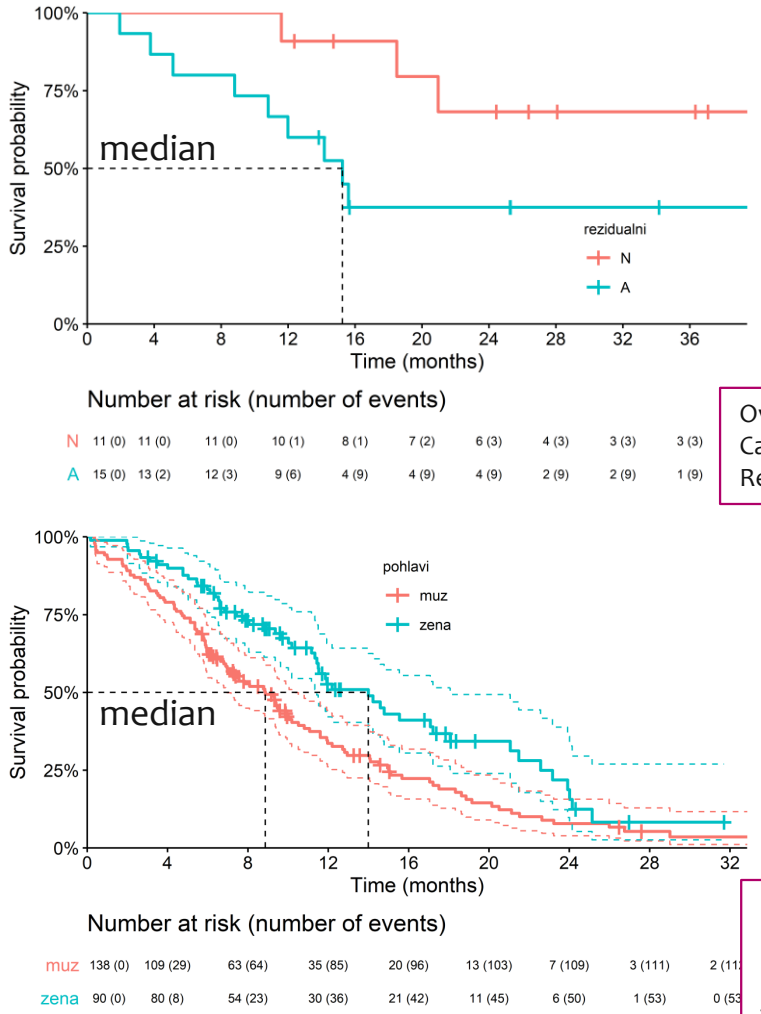
T_1, date of diagnosis; T_2, final recorded date; T_3, data cut-off date

Reporting follow-up in survival analyses: informative or not? Erasmus School of Health Policy & Management

Basic Quantities of Survival Time



Wikipedie



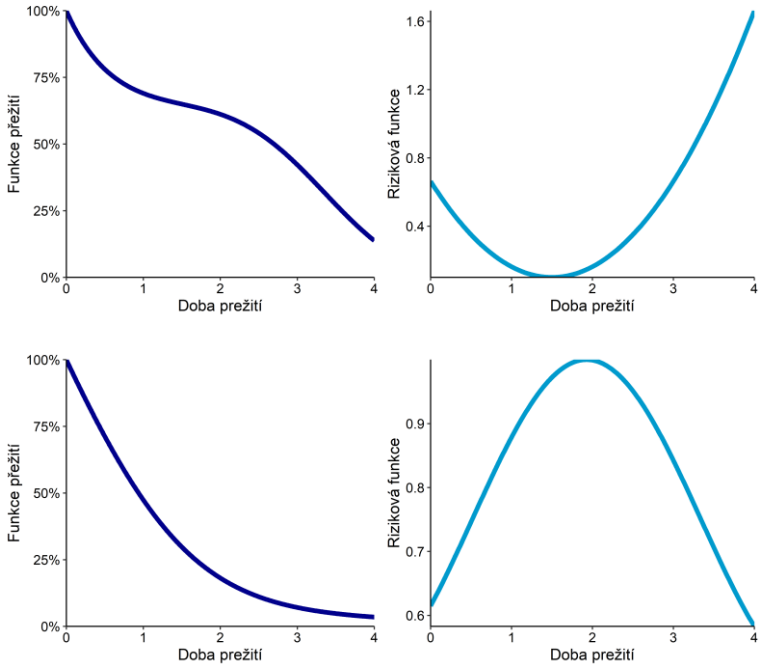
Ovarian Carcinoma, Cancer Treatment Reports 1979

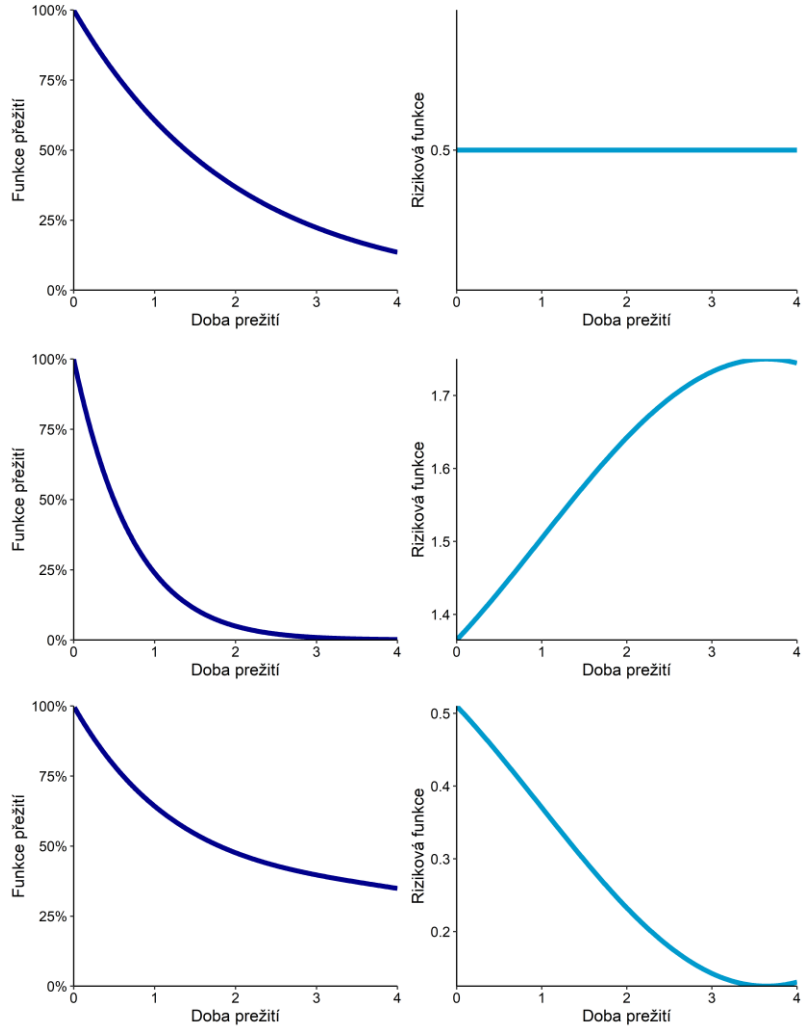
Lung Cancer, North Central Cancer Treatment Group. Journal of Clinical Oncology 1994

Basic Quantities of Survival Time

Hazard function

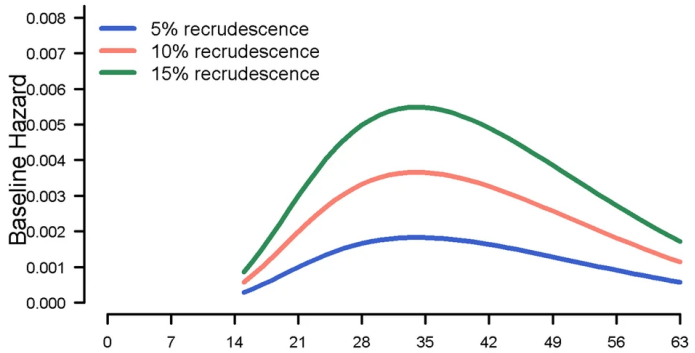
The intensity of the occurrence of the event at time, given that the individual survived that time

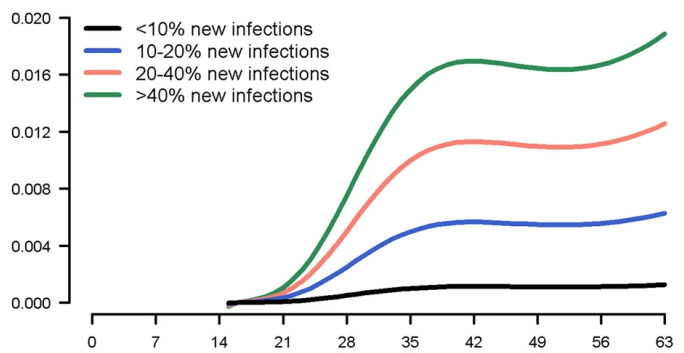


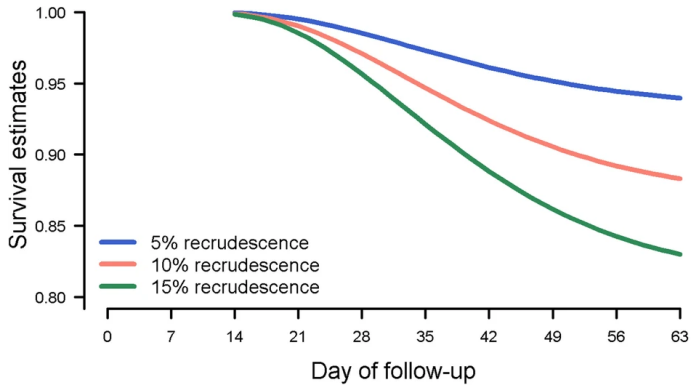


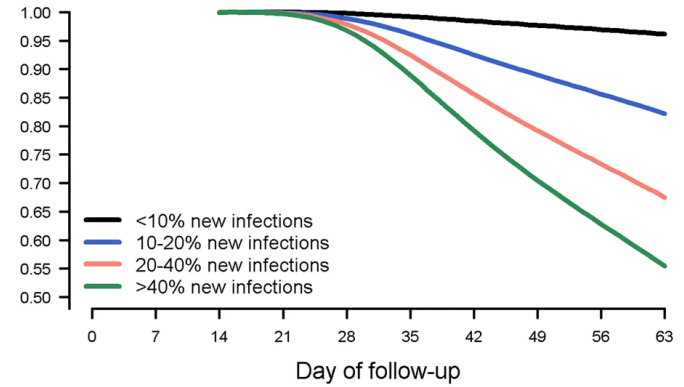
Basic Quantities of Survival Time

Proportional Hazard Function









BMC Medical Research Methodology 2019
Evaluating antimalarial efficacy in single-armed and comparative drug trials using competing risk survival analysis: a simulation study

Hazard Ratio (HR)

- Ratio of hazard function of two different groups
- Proportional hazard function: HR is time-independent

Qualitative data

- One category is reference
- HR represents how many times higher/lower the risk of occurrence of event is for a given category compared to reference category
- HR=1 means that the risk is equal in both groups

Male vs Female (ref)

$HR_{male}=2.5$

- Event risk is 2.5×higher for males than for females (higher by 150%)
- $HR_{female}=1/2.5=0.4$, i.e., risk for females is 0.4×lower than for males (lower by 60%)

Arm A vs Arm B vs Placebo (ref)

$HR_A=1.2$, $HR_B=0.6$

- Treatment A has 1.2×higher event risk than placebo (higher by 20%)
- Treatment B has 0.6×lower event risk than placebo (lower by 40 %)

Hazard Ratio (HR)

- Ratio of hazard function of two different groups
- Proportional hazard function: HR is time-independent

Quantitative data

- HR represents how many times is changed event risk when the predictor value increases by one
- HR=1 means that the risk is unchanged, survival is the same

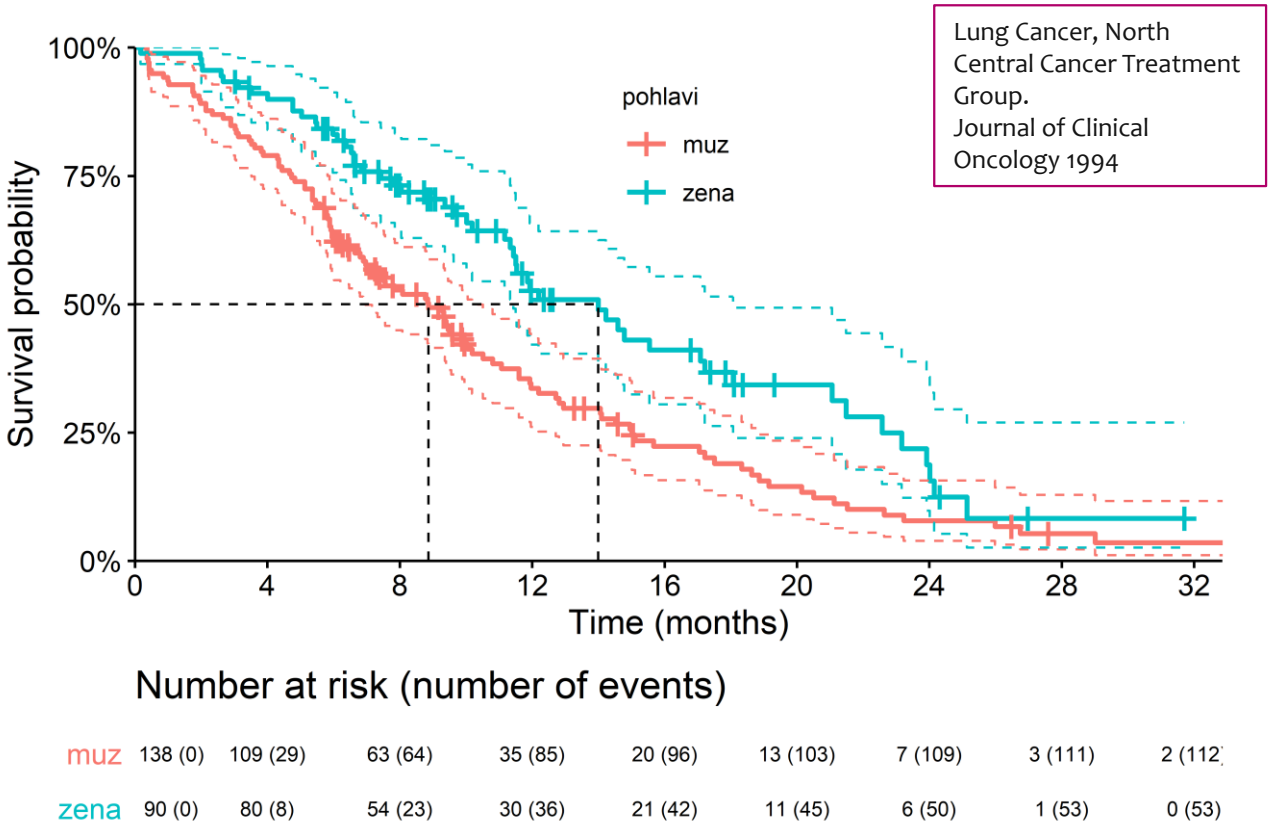
Age (years) HR=1.05

- Event risk increases 1.05× with each one year of age (by 5%)
- E.g. 53-year-old patient has the event risk by 5% higher than 52-year-old
- 60-year-old patient vs 60-year-old patient: $1.05^{10}=1.63$, i.e. 60-year-old patient has the event risk by 63% higher than 50-year-old

Hazard Ratio (HR) Interpretation

Hazard ratio	Event risk	Survival
HR=1	Equal	Equal
HR<1	Lower	Longer
HR>1	Higher	Shorter

Comparison of Survival Function

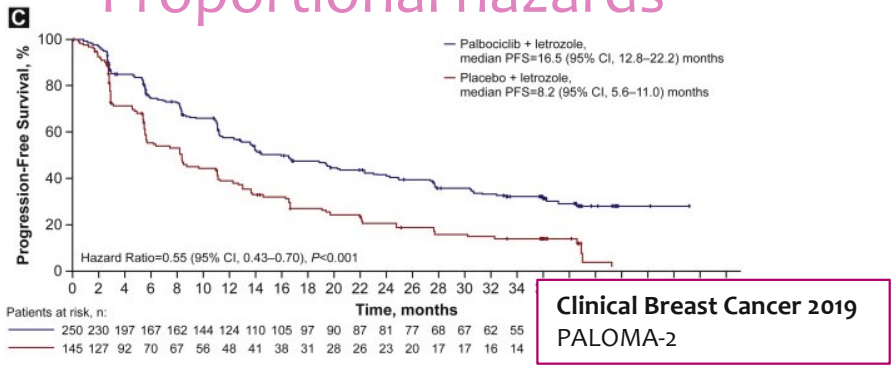


Lung Cancer, North Central Cancer Treatment Group. Journal of Clinical Oncology 1994

- Various statistical tests
- If proportionality hazards assumption is met, **HR=1** for equal survival function

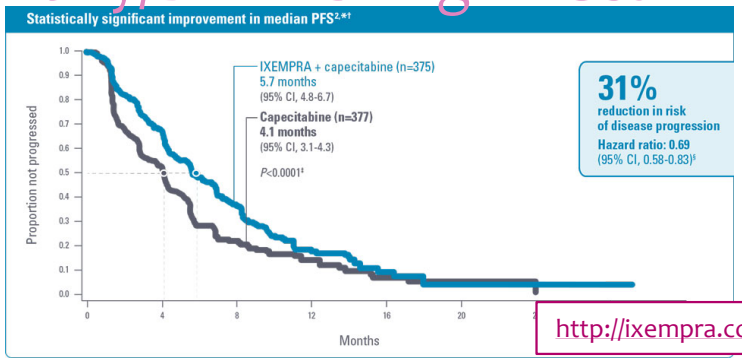
Comparison of Survival Function

Proportional hazards



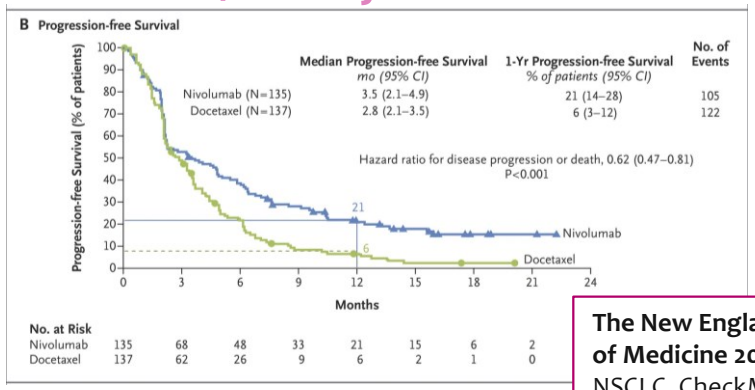
Clinical Breast Cancer 2019
PALOMA-2

Early/Diminishing Effect



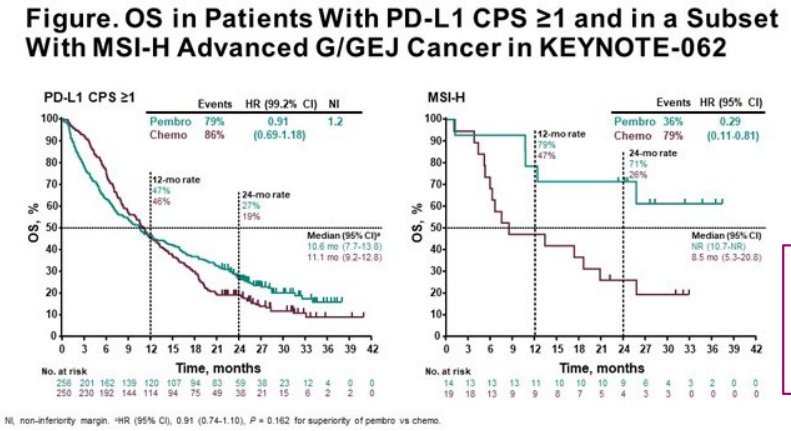
http://ixempra.com/hcp/combo

Late/Delayed Effect



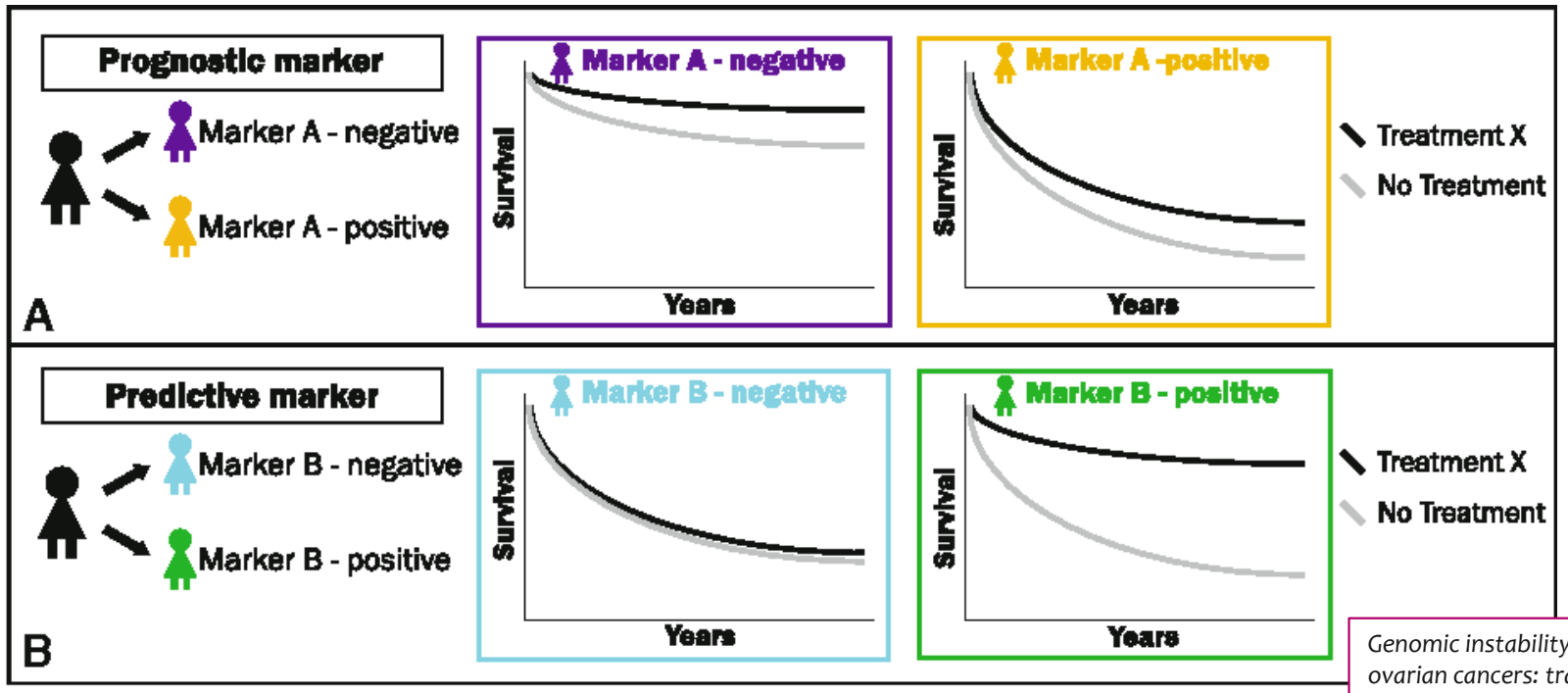
The New England Journal of Medicine 2015
NSCLC, CheckMate 017

Crossing Hazards



Adenocarcinoma, KEYNOTE-062 study presented at the **ESMO Congress 2019**

Prognostic and Predictive Factors



Genomic instability in breast and ovarian cancers: translation into clinical predictive biomarkers, Cellular and Molecular Life Sciences, 2012

Prognostic:
determines outcome regardless of treatment
Predictive:
determines the success of the treatment

Take-Home Message

- Choosing the right study endpoints is one of the key steps in clinical trial design
- The primary objective/endpoint affects the study design
- Endpoints are measurable outcomes used to address the objectives of a clinical trial, such as survival, decreased pain, or the absence of disease
- Study endpoints can be divided into direct and surrogate, must also be taken into account when interpreting the results of the study
- For time-to-event data specification and afterward interpretation, it is necessary to specify the event of interest and the beginning of follow-up
- HR and OR are used to express the strength of the intervention effect

Thank you for your attention