



University of  
Zurich <sup>UZH</sup>

# BIO392

# Bioinformatics of Genome Variations

Survival | Classifications

Michael Baudis **UZH SIB**  
Computational Oncogenomics

# BIO392 HS 2023

## Github Activity

September 3, 2023 – October 3, 2023

Period: 1 month ▾

### Overview

2 Active pull requests

0 Active issues

2

Merged pull requests

0

Open pull requests

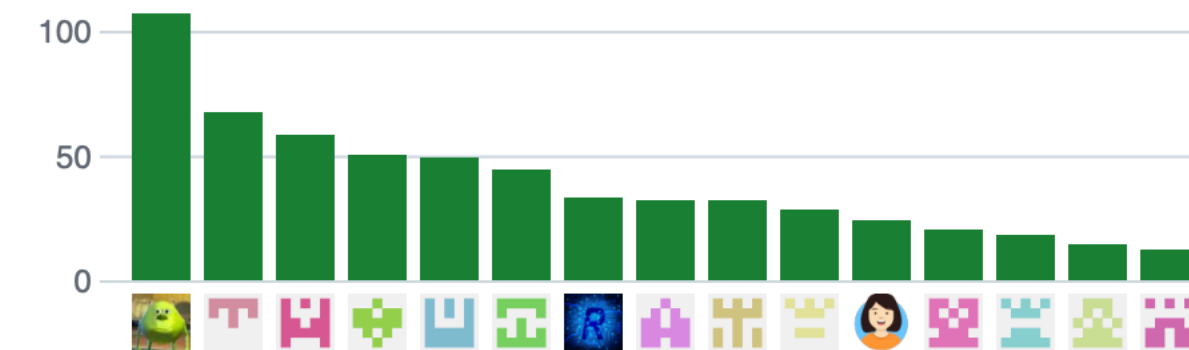
0

Closed issues

0

New issues

Excluding merges, **19 authors** have pushed **610 commits** to master and **611 commits** to all branches. On master, **299 files** have changed and there have been **10,596 additions** and **8,144 deletions**.



2 Pull requests merged by 1 person

### Simplify wiggle

#220 merged 2 weeks ago

### Update 2023 slides, copy exercises from 2022

#219 merged 2 weeks ago

# BIO392 HS 2022

## Github Activity

September 3, 2023 – October 3, 2023

Period: 1 month ▾

**Overview**

2 Active pull requests      0 Active issues

2 Merged pull requests      0 Open pull requests

Excluding merges, **19 authors** have pushed **610 commits** to master and **611 commits** to all branches. On master, **299 files** have changed and there have been **10,596 additions** and **8,144 deletions**.

Eric-Thiele  
Committed to this repository in the past week  
Member of [compbiozurich/bio392-2023](#)  
Member of [CompbioZurich](#)  
Joined GitHub this month

2 Pull requests merged by 1 person

**Simplify wiggle**  
#220 merged 2 weeks ago

**Update 2023 slides, copy exercises from 2022**  
#219 merged 2 weeks ago

# Survival

**Kaplan-Meier Analysis of Survival Based on Conditional Probabilities**

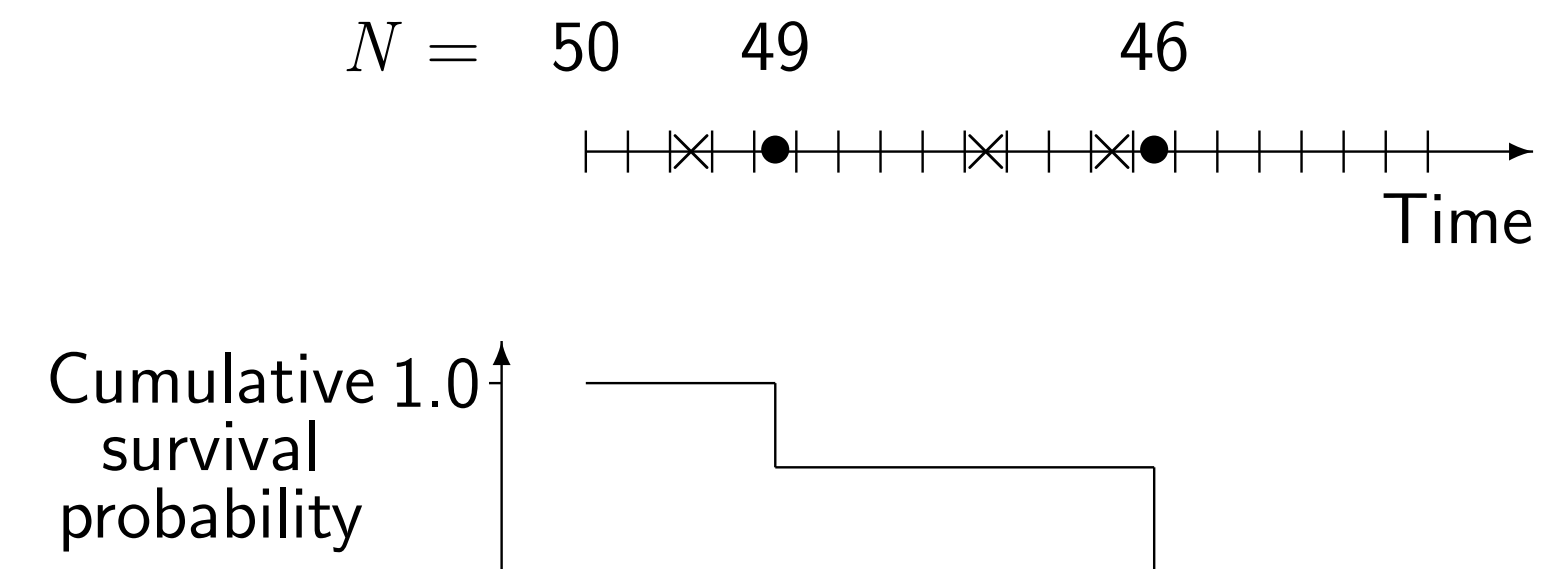
## The Kaplan-Meier Method

- ▶ The most common method of estimating the survival function.
- ▶ A non-parametric method.
- ▶ Divides time into small intervals where the intervals are defined by the unique times of failure (death).
- ▶ Based on conditional probabilities as we are interested in the probability a subject surviving the next time interval given that they have survived so far.

Kaplan-Meier estimators (km-na)

## Kaplan–Meier method illustrated

(● = failure and × = censored):



- ▶ Steps caused by multiplying by  $(1 - 1/49)$  and  $(1 - 1/46)$  respectively
- ▶ Late entry can also be dealt with

Kaplan-Meier estimators (km-na)

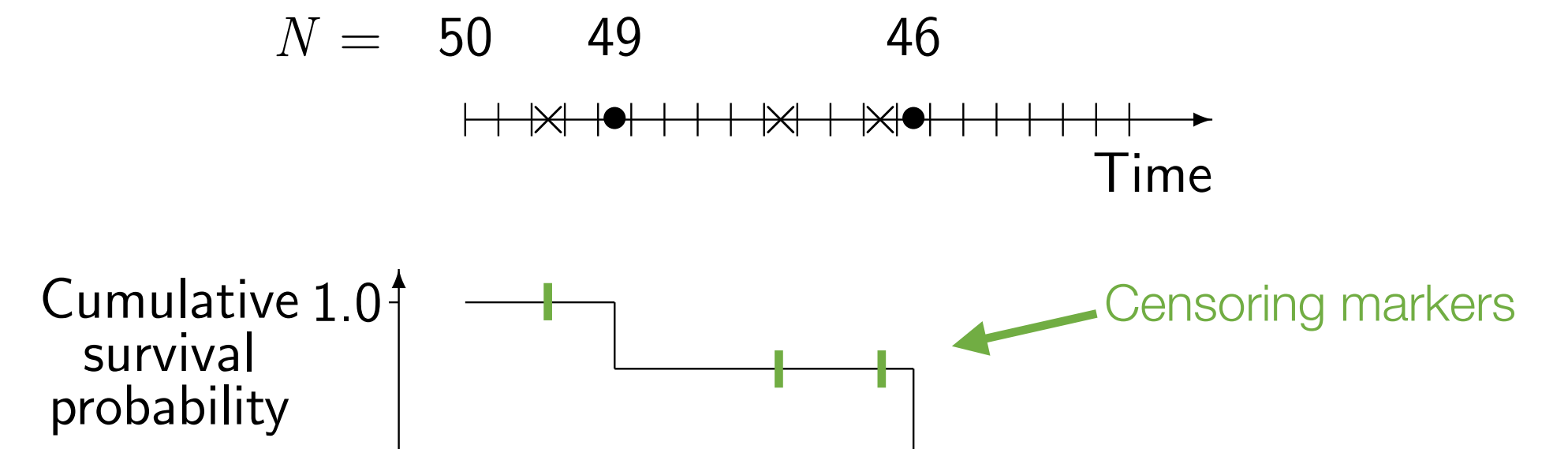
## The Kaplan-Meier Method

- ▶ The most common method of estimating the survival function.
- ▶ A non-parametric method.
- ▶ Divides time into small intervals where the intervals are defined by the unique times of failure (death).
- ▶ Based on conditional probabilities as we are interested in the probability a subject surviving the next time interval given that they have survived so far.

Kaplan-Meier estimators (km-na)

## Kaplan–Meier method illustrated

(● = failure and × = censored):

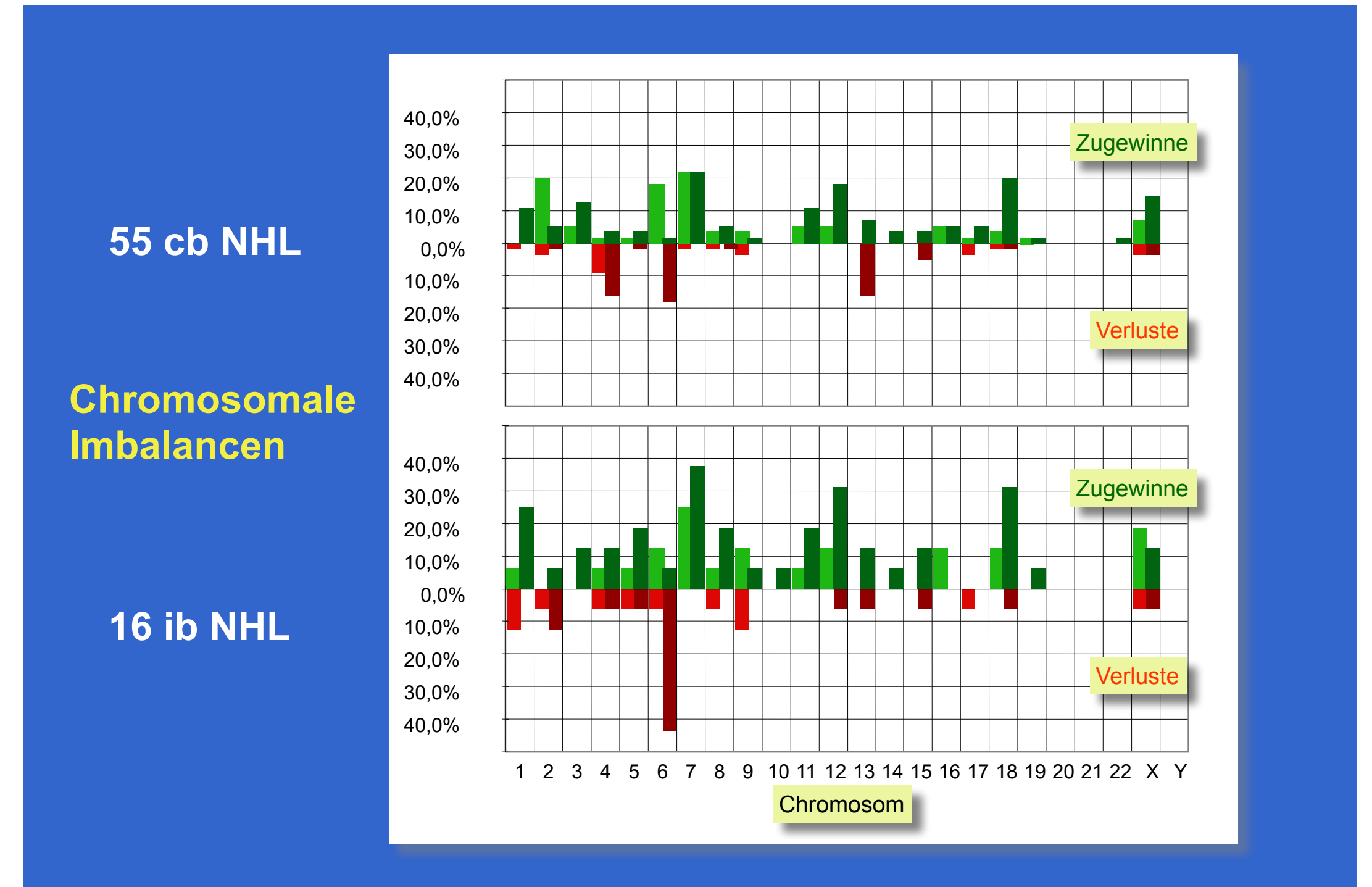
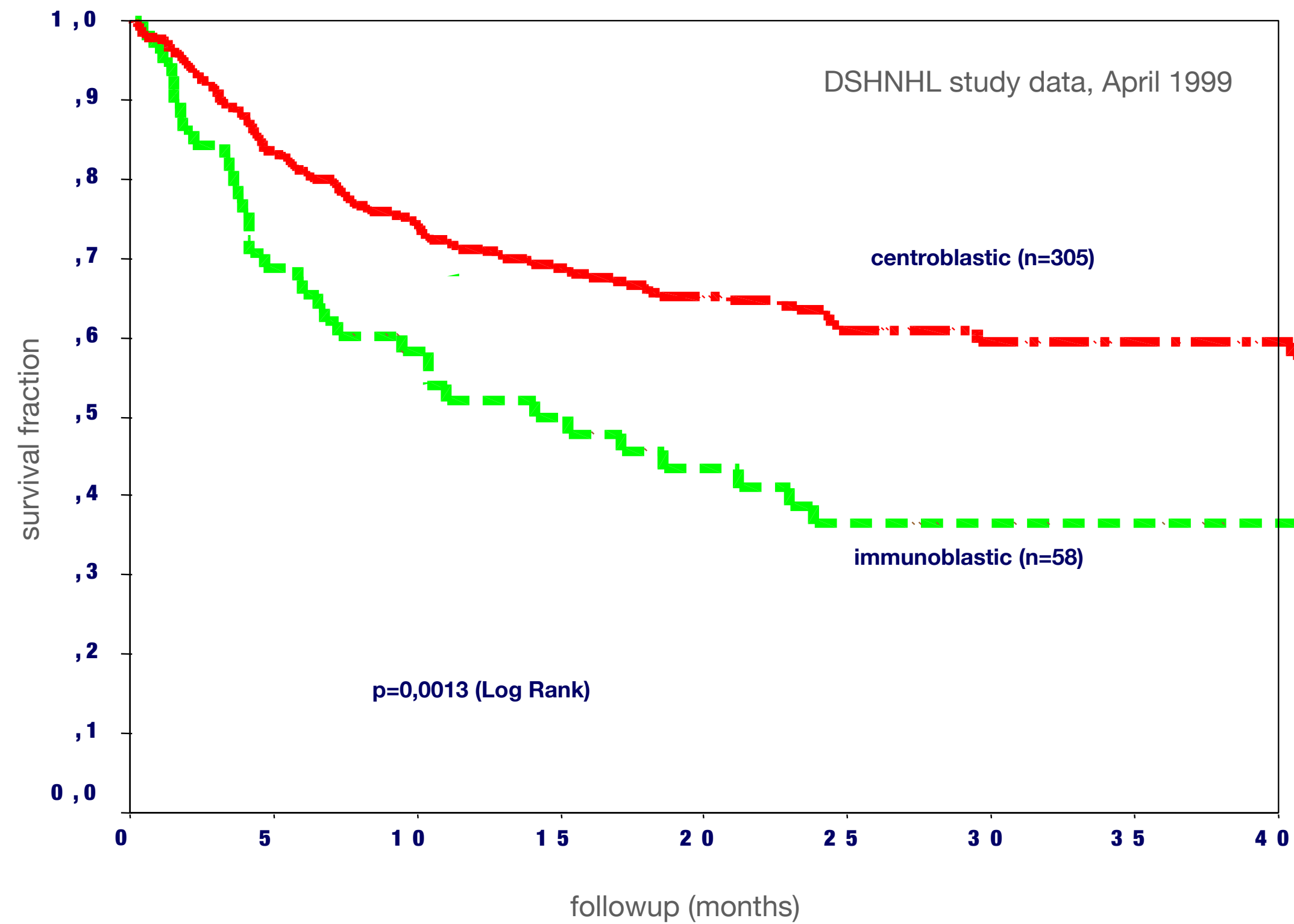


- ▶ Steps caused by multiplying by  $(1 - 1/49)$  and  $(1 - 1/46)$  respectively
- ▶ Late entry can also be dealt with

Kaplan-Meier estimators (km-na)

# Cancer CNVs | Diagnostics | Prognosis

## Single-study CNV frequencies correspond to diagnostic subsets



# Kaplan-Meier Plots to Visualize Differential Risk

## Multi-parametric "risk scores" in CLL Prognosis

Leukemia (2020) 34:1038–1051  
<https://doi.org/10.1038/s41375-020-0727-y>

### ARTICLE

Chronic lymphocytic leukemia

#### Prognostic model for newly diagnosed CLL patients in Binet stage A: results of the multicenter, prospective CLL1 trial of the German CLL study group

Manuela A. Hoehstetter<sup>1</sup> · Raymonde Busch<sup>2</sup> · Barbara Eichhorst<sup>3</sup> · Andreas Bühler<sup>4</sup> · Dirk Winkler<sup>4</sup> · Jasmin Bahlo<sup>3</sup> · Sandra Robrecht<sup>3</sup> · Michael J. Eckart<sup>5</sup> · Ursula Vehling-Kaiser<sup>6</sup> · Georg Jacobs<sup>7</sup> · Ulrich Jäger<sup>8</sup> · Hans Jürgen Hurtz<sup>9</sup> · Georg Hopfinger<sup>10</sup> · Frank Hartmann<sup>11</sup> · Harald Fuss<sup>12</sup> · Wolfgang Abenhardt<sup>13</sup> · Ilona Blau<sup>14</sup> · Werner Freier<sup>15</sup> · Lothar Müller<sup>16</sup> · Maria Goebeler<sup>17</sup> · Clemens Wendtner<sup>13</sup> · Kirsten Fischer<sup>3</sup> · Carmen D. Herling<sup>3</sup> · Michael Starck<sup>1</sup> · Martin Bentz<sup>18</sup> · Bertold Emmerich<sup>19</sup> · Hartmut Döhner<sup>20</sup> · Stephan Stilgenbauer<sup>20</sup> · Michael Hallek<sup>3</sup>

**Table 2a** Results of the Cox's regression for OS and TTFT in CLL patients in whom all 30 baseline parameters were available.

Univariate comparison	Hazard ratio [HR]	95% Confidence Interval		P value
		Lower	Upper	
<b>COX regression OS</b>				
Cytogenetic Hierarchical Type				
del(17p) vs. not del(17p)/del(11q)	3.8	2.1	7.1	<0.001
del(11q) vs. not del(17p)/del(11q)	2.0	1.2	3.5	0.008
LDT				
<12 months vs. ≥12 months	1.9	1.3	2.8	0.001
Age, years				
>60 vs. ≤60	1.8	1.2	2.7	0.002
B2M, mg/dL				
>3.5 vs. ≤3.5	2.0	1.2	3.1	0.004
IGHV mutational status				
Unmutated vs. mutated	2.4	1.6	3.6	<0.001
<b>COX regression TTFT</b>				
Cytogenetic Hierarchical Type				
del(17p) vs. not del(17p)/del(11q)	2.2	1.2	4.1	0.009
del(11q) vs. not del(17p)/del(11q)	2.0	1.3	3.0	0.001
LDT vs.	2.3	1.7	3.1	<0.001
Age, years				
>60 vs. ≤60	1.3	1.0	1.7	0.037
B2M, mg/dL				
>3.5 vs. ≤3.5	1.5	1.0	2.3	0.049
IGHV mutational status				
Unmutated vs. mutated	4.4	3.2	5.9	<0.001

**Table 2b** Allocation of risk score points to the distinctive factors of the CLL1-PM.

Characteristics	HR (95% CI)	P	Allocated risk score points
Del(17p)	3.8 (2.1–7.1)	<0.001	3.5
Unmutated IGHV	2.4 (1.6–3.6)	<0.001	2.5
Del(11q)	2.0 (1.2–3.5)	0.008	2.5
Beta2-MG >3.5 mg/L	2.0 (1.2–3.1)	0.004	2.5
LDT<12 months	1.9 (1.3–2.8)	0.001	1.5
Age >60 years	1.8 (1.2–2.7)	0.002	1.5

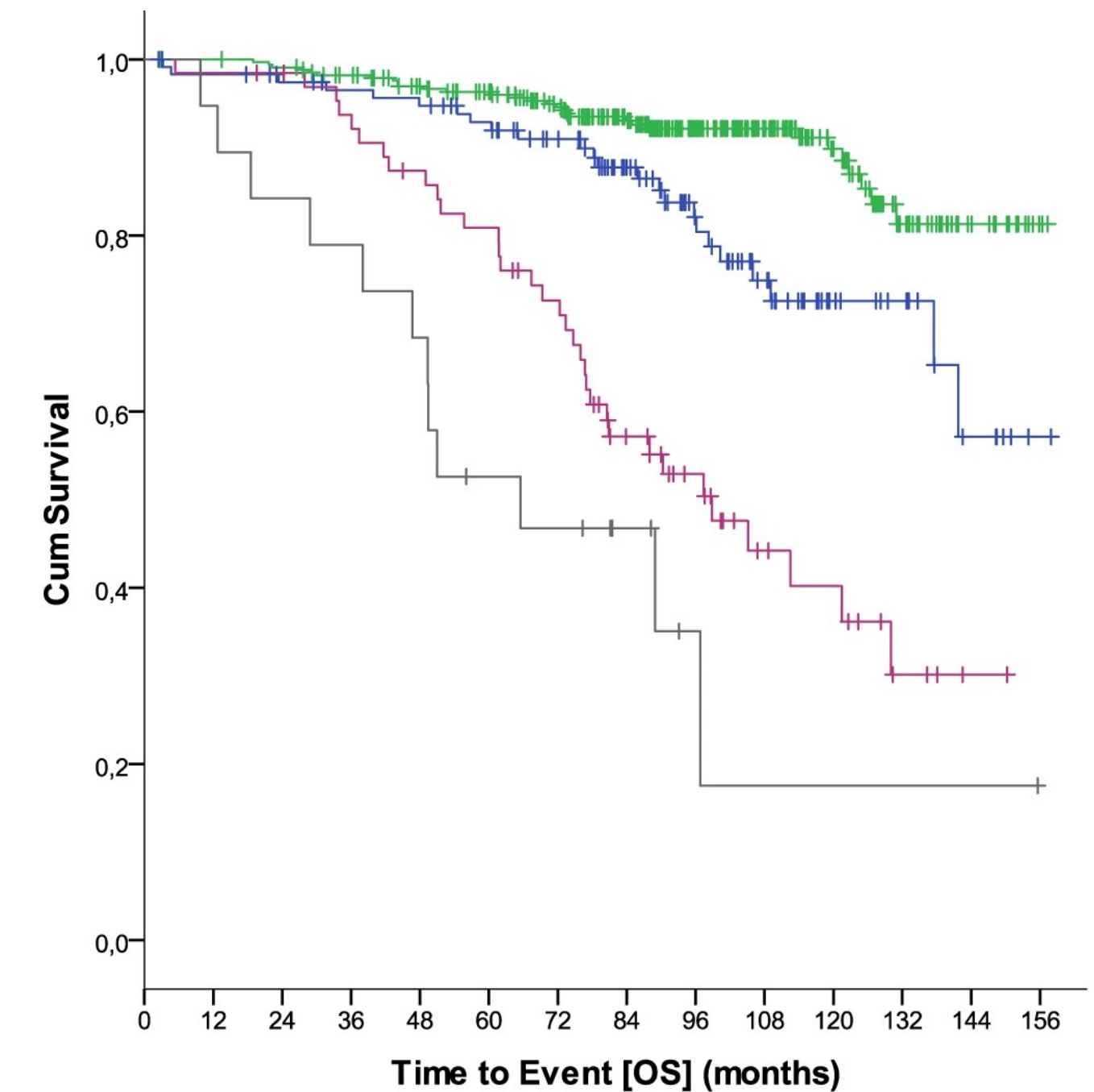
The assigned risk score points derived from the HR for OS of the individual factors.

**Table 2c** Patients and risk groups according to the CLL1 Prognostic Model (CLL1-PM). Patients and risk groups according to the CLL-IPI.

	Index score	Patients N (%)
Risk Groups according to the CLL1-PM		
Very low	0.0–1.5	336 (62.3)
Low	2.0–4.0	119 (22.1)
High	4.5–6.5	65 (12.1)
Very high	7.0–14.0	19 (3.5)
Risk Groups according to the CLL-IPI		
Low	0–1	360 (66.8)
Intermediate	2–3	141 (26.2)
High	4–6	33 (6.1)
Very high	7–10	5 (0.9)

OS overall survival, HR hazard ratio, Beta2-MG beta-2 microglobulin, IGHV immunoglobulin heavy-chain genes, LDT lymphocyte doubling time, TTFT time-to-first treatment.

- "a novel prognostic model (CLL1-PM) developed to identify risk groups, separating patients with favorable from others with dismal prognosis"
- " findings would be useful to effectively stratify Binet stage A patients, particularly within the scope of clinical trials evaluating novel agents"



**P < 0.001**

Number at risk	0	12	24	36	48	60	72	84	96	108	120	132	144	156
Very low	336	335	331	322	306	294	262	215	160	113	68	33	15	2
Low	119	115	111	108	106	100	89	71	49	34	19	14	6	1
High	65	64	63	59	54	50	43	29	21	12	10	4	1	0
Very high	19	18	16	15	13	9	8	5	2	1	1	1	1	0

**Discrimination:** C-statistics, C = 0.739 (95% CI, 0.686– 0.790)  
**AIC=445**

Overall survival according to the CLL1-PM risk groups. The full analysis dataset is comprised of the dataset of 539 patients.



# Kaplan-Meier Plots to Visualize Differential Risk

## Multi-parametric "risk scores" in CLL Prognosis

Leukemia (2020) 34:1038–1051  
<https://doi.org/10.1038/s41375-020-0727-y>

### ARTICLE

Chronic lymphocytic leukemia

### Prognostic model for newly diagnosed CLL patients in Binet stage A: results of the multicenter, prospective CLL1 trial of the German CLL study group

Manuela A. Hoehstetter<sup>1</sup> · Raymonde Busch<sup>2</sup> · Barbara Eichhorst<sup>3</sup> · Andreas Bühler<sup>4</sup> · Dirk Winkler<sup>4</sup> · Jasmin Bahlo<sup>3</sup> · Sandra Robrecht<sup>3</sup> · Michael J. Eckart<sup>5</sup> · Ursula Vehling-Kaiser<sup>6</sup> · Georg Jacobs<sup>7</sup> · Ulrich Jäger<sup>8</sup> · Hans Jürgen Hurtz<sup>9</sup> · Georg Hopfinger<sup>10</sup> · Frank Hartmann<sup>11</sup> · Harald Fuss<sup>12</sup> · Wolfgang Abenhardt<sup>13</sup> · Ilona Blau<sup>14</sup> · Werner Freier<sup>15</sup> · Lothar Müller<sup>16</sup> · Maria Goebeler<sup>17</sup> · Clemens Wendtner<sup>13</sup> · Kirsten Fischer<sup>3</sup> · Carmen D. Herling<sup>3</sup> · Michael Starck<sup>1</sup> · Martin Bentz<sup>18</sup> · Bertold Emmerich<sup>19</sup> · Hartmut Döhner<sup>20</sup> · Stephan Stilgenbauer<sup>20</sup> · Michael Hallek<sup>3</sup>

**Table 2a** Results of the Cox's regression for OS and TTFT in CLL patients in whom all 30 baseline parameters were available.

Univariate comparison	Hazard ratio [HR]	95% Confidence Interval		P value
		Lower	Upper	
<b>COX regression OS</b>				
Cytogenetic Hierarchical Type				
del(17p) vs. not del(17p)/del(11q)	3.8	2.1	7.1	<0.001
del(11q) vs. not del(17p)/del(11q)	2.0	1.2	3.5	0.008
LDT				
<12 months vs. ≥12 months	1.9	1.3	2.8	0.001
Age, years				
>60 vs. ≤60	1.8	1.2	2.7	0.002
B2M, mg/dL				
>3.5 vs. ≤3.5	2.0	1.2	3.1	0.004
IGHV mutational status				
Unmutated vs. mutated	2.4	1.6	3.6	<0.001
<b>COX regression TTFT</b>				
Cytogenetic Hierarchical Type				
del(17p) vs. not del(17p)/del(11q)	2.2	1.2	4.1	0.009
del(11q) vs. not del(17p)/del(11q)	2.0	1.3	3.0	0.001
LDT vs.	2.3	1.7	3.1	<0.001
Age, years				
>60 vs. ≤60	1.3	1.0	1.7	0.037
B2M, mg/dL				
>3.5 vs. ≤3.5	1.5	1.0	2.3	0.049
IGHV mutational status				
Unmutated vs. mutated	4.4	3.2	5.9	<0.001

**Table 2b** Allocation of risk score points to the distinctive factors of the CLL1-PM.

Characterization	HR (95% CI)	P	Allocated risk score points
Del(17p)	3.8 (2.1–7.1)	<0.001	3.5
Del(11q)	2.0 (1.2–3.5)	0.008	2.5
Beta2-MG >3.5 mg/L	2.0 (1.2–3.1)	0.004	2.5
LDT <12 months	1.9 (1.3–2.8)	0.001	1.5
Age >60 years	1.8 (1.2–2.7)	0.002	1.5

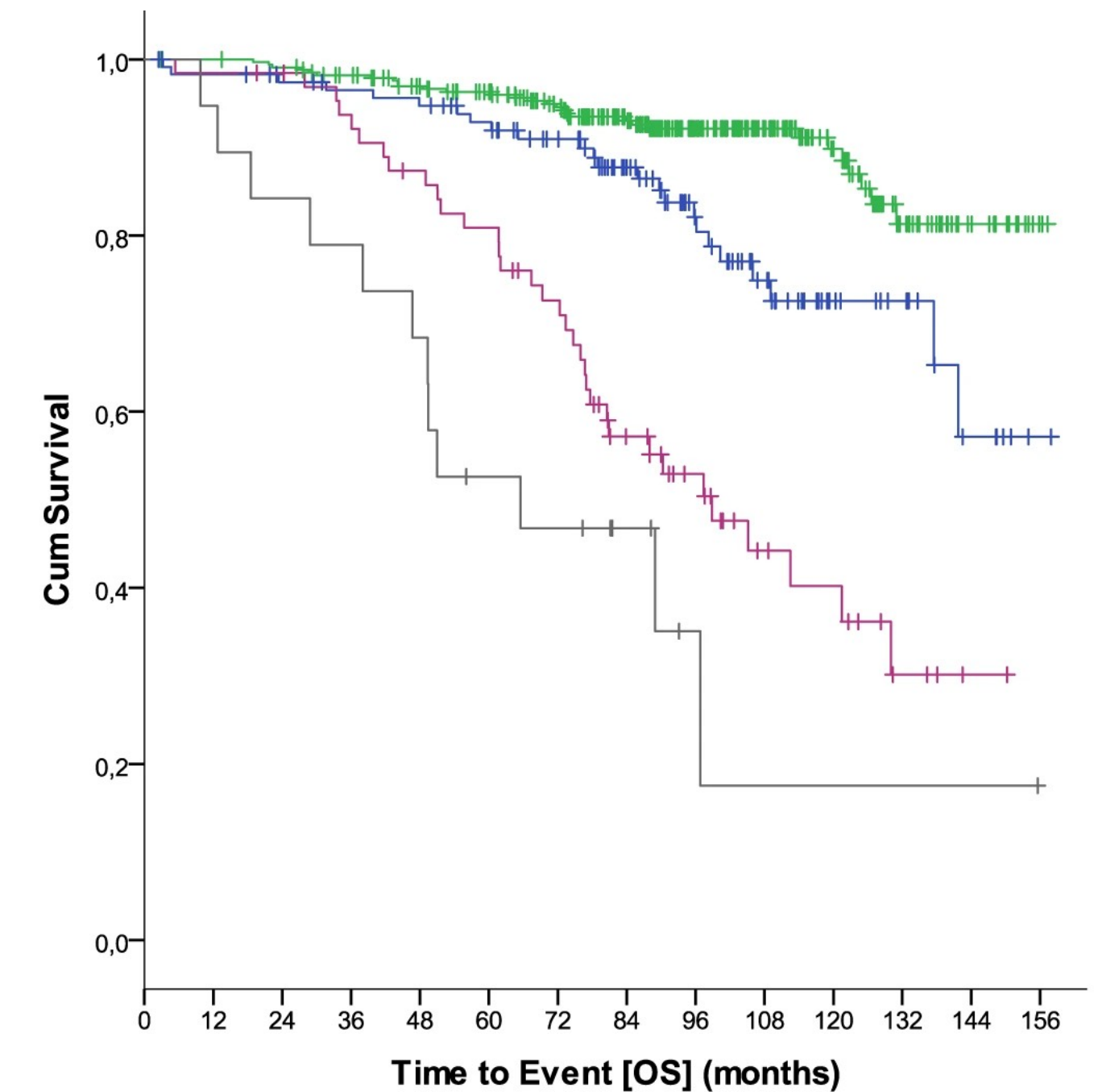
The assigned risk score points derived from the HR for OS of the individual factors.

**Table 2c** Patients and risk groups according to the CLL1 Prognostic Model (CLL1-PM). Patients and risk groups according to the CLL-IPI.

	Index score	Patients N (%)
Risk Groups according to the CLL1-PM		
Very low	0–1.5	336 (62.3)
Low	2.0–4.0	119 (22.1)
High	4.5–6.5	65 (12.1)
Very high	7.0–14.0	19 (3.5)
Risk Groups according to the CLL-IPI		
Low	0–1	360 (66.8)
Intermediate	2–3	141 (26.2)
High	4–6	33 (6.1)
Very high	7–10	5 (0.9)

OS overall survival, HR hazard ratio, Beta2-MG beta-2 microglobulin, IGHV immunoglobulin heavy-chain genes, LDT lymphocyte doubling time, TTFT time-to-first treatment.

- "a novel prognostic model (CLL1-PM) developed to identify risk groups, separating patients with favorable from others with dismal prognosis"
- "findings would be useful to effectively stratify Binet stage A patients, particularly within the scope of clinical trials evaluating novel agents"



**P < 0.001**

Number at risk	0	12	24	36	48	60	72	84	96	108	120	132	144	156
Very low	336	335	331	322	306	294	262	215	160	113	68	33	15	2
Low	119	115	111	108	106	100	89	71	49	34	19	14	6	1
High	65	64	63	59	54	50	43	29	21	12	10	4	1	0
Very high	19	18	16	15	13	9	8	5	2	1	1	1	1	0

**Discrimination:** C-statistics, C = 0.739 (95% CI, 0.686– 0.790)  
**AIC=445**

Overall survival according to the CLL1-PM risk groups. The full analysis dataset is comprised of the dataset of 539 patients.

# Cancer Classifications & Parameters

NCIt | ICD-O / WHO | TNM

# ICD-O 3

## WHO International Classification of Diseases for Oncology, 3rd Edition (ICD-O-3)

- used in cancer registries for coding the site (topography) and the histology (morphology) of neoplasms, usually obtained from a pathology report
- mix of "biology" (i.e. tumor morphology) and "clinical" (i.e. tumor site)

➔ 2 codes per cancer

- ▶ "Adenocarcinoma" of the "Sigmoid colon"

**8140/3**

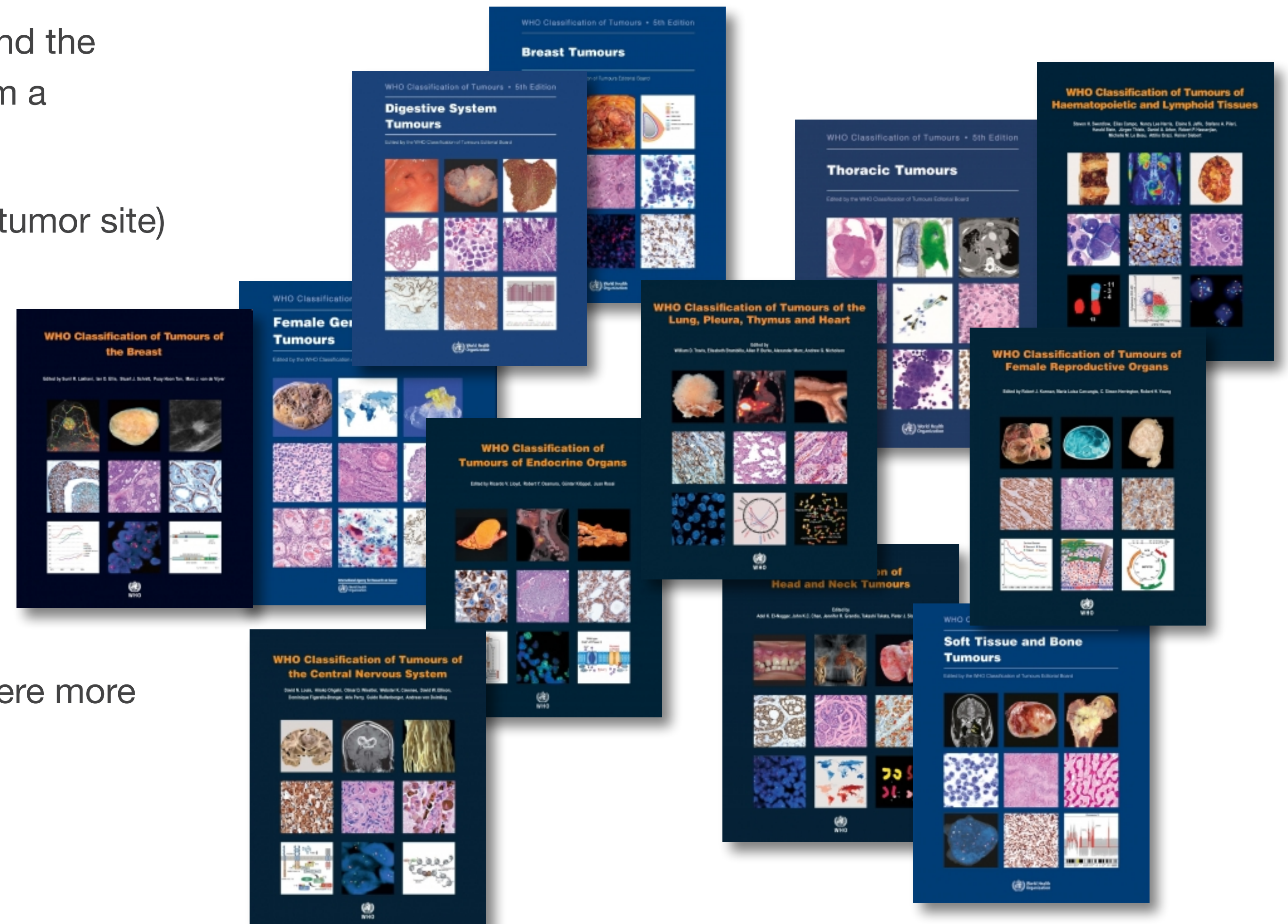
**C18.7**

- ▶ "Retinoblastoma" of the "Retina"

**9510/3**

**C69.2**

- widely accepted by pathologists but limited clinical use (there more ICD-10 or SNOMED)
- no ontology & not (truly) hierarchical
- many entities difficult to remap if using only single code



# NCIt

## Neoplasm Classifications in the NCI Thesaurus

- NCI's core reference terminology and biomedical ontology are collected in the NCI Thesaurus (NCIt)
- individual codes for site-specific occurrences of "biological" diagnoses

### 1 code per cancer

- ▶ **NCIT:C43584** - Rectosigmoid Adenocarcinoma
- ▶ **NCIT:C7541** - Retinoblastoma
- truly hierarchical ontology
- hierarchical system empowers "logical OR" queries
- terms can have multiple occurrences in diagnostic tree
- assignment of code to different groupings allows soft aggregation (e.g. a type of colorectal adenocarcinoma with all colon tumors or with all adenocarcinomas)

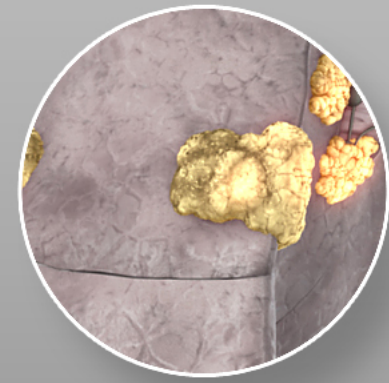


- ▼ **NCIT:C3262**: Neoplasm (116013 samples)
  - ▼ **NCIT:C3263**: Neoplasm by Site (110893 samples)
    - ▶ **NCIT:C156482**: Genitourinary System Neoplasm (16534 samples)
    - ▶ **NCIT:C2910**: Breast Neoplasm (15957 samples)
    - ▶ **NCIT:C3010**: Endocrine Neoplasm (3521 samples)
    - ▶ **NCIT:C3030**: Eye Neoplasm (280 samples)
  - ▼ **NCIT:C3052**: Digestive System Neoplasm (15289 samples)
    - ▶ **NCIT:C172852**: Digestive System Soft Tissue Neoplasm (99 samples)
    - ▶ **NCIT:C27721**: Digestive System Neuroendocrine Neoplasm (202 samples)
    - ▶ **NCIT:C2877**: Anal Neoplasm (61 samples)
    - ▶ **NCIT:C3028**: Esophageal Neoplasm (1865 samples)
  - ▼ **NCIT:C3141**: Intestinal Neoplasm (5723 samples)
    - ▼ **NCIT:C2956**: Colorectal Neoplasm (5579 samples)
      - ▶ **NCIT:C2953**: Colon Neoplasm (4666 samples)
      - ▶ **NCIT:C3350**: Rectal Neoplasm (527 samples)
      - ▶ **NCIT:C4610**: Benign Colorectal Neoplasm (181 samples)
    - ▼ **NCIT:C4877**: Rectosigmoid Neoplasm (240 samples)
      - ▼ **NCIT:C7420**: Malignant Rectosigmoid Neoplasm (240 samples)
        - ▼ **NCIT:C7421**: Rectosigmoid Carcinoma (240 samples)
          - ▼ **NCIT:C43584**: Rectosigmoid Adenocarcinoma (240 samples)
            - ▶ **NCIT:C43592**: Rectosigmoid Mucinous Adenoca... (18 samples)
- ▶ **NCIT:C4978**: Malignant Colorectal Neoplasm (5398 samples)
- ▶ **NCIT:C96152**: Colorectal Neuroendocrine Neoplasm (11 samples)
- ▶ **NCIT:C4432**: Small Intestinal Neoplasm (66 samples)

# TNM

## A Classification for Clinical Cancer Stage Parameters

- most widely used cancer staging system
- **T** refers to the size and extent of the main tumor
- **N** refers to the the number / location of nearby lymph nodes that have cancer infiltration
- **M** refers to whether the cancer has metastasized
- not used for leukemias / lymphomas
  - Binet and Rai in CLL
  - proportion of blasts in bone marrow or blood in leukemias
  - Lugano classification in lymphomas
- other disease specific staging systems may (co-) exist
  - e.g. a stage II breast cancer is determined by size & nodal involvement

Source: <https://www.cancer.gov/about-cancer/diagnosis-staging/staging>

 <p><b>T</b>umor size</p> <p><b>T1</b> – A tumor is less than 3 cm (1 ½ inches) in size <b>T2</b> – The tumor is greater than 3 cm <b>T3</b> – The tumor can be any size, but is near the airway or has spread to local areas such as the chest wall or diaphragm <b>T4</b> – The tumor is any size, but is located in the airway, or has invaded local structures such as the heart or the esophagus.</p>	 <p>Lymph <b>N</b>odes</p> <p><b>N0</b> – No lymph nodes are affected <b>N1</b> – The tumor has spread to nearby nodes on the same side of the body <b>N2</b> – The tumor has spread to nodes further away but on the same side of the body <b>N3</b> – Cancer cells are present in lymph nodes on the other side of the chest from the tumor, or in nodes near the collarbone or neck muscles</p>	 <p><b>M</b>etastases</p> <p><b>M0</b> – No metastases are present <b>M1</b> – The tumor has spread (metastasized) to other regions of the body or the other lung</p>
---	---	--

Source: [www.scientificanimations.com](http://www.scientificanimations.com)

# TNM

## A Classification for Clinical Cancer Stage Parameters

- most widely used cancer staging system
- **T** refers to the size and extent of the main tumor
- **N** refers to the the number / location of nearby lymph nodes that have cancer infiltration
- **M** refers to whether the cancer has metastasized
- not used for leukemias / lymphomas
  - ▶ Binet and Rai in CLL
  - ▶ proportion of blasts in bone marrow or blood in leukemias
  - ▶ Lugano classification in lymphomas
- other disease specific staging systems may (co-) exist
  - ▶ e.g. a stage II lung cancer is determined by size & nodal involvement

Source: <https://www.cancer.gov/about-cancer/diagnosis-staging/staging>

## TNM STAGING OF LUNG CANCER - 8<sup>th</sup> EDITION

<b>Stage IV B (Any T, Any N, M1c)</b>		<b>M1c</b>	Multiple extrathoracic metastases (in one or more organs)
<b>Stage IV A (Any T, Any N, M1a/b)</b>		<b>M1b</b>	Single extrathoracic metastasis (including non-regional lymph nodes)
		<b>M1a</b>	Satellite (separate) tumor nodule(s) in contralateral lobe or Pleural or pericardial nodules or malignant effusion
<b>Distant metastasis</b>		<b>M1</b>	<b>DISTANT METASTASIS (M)</b>
<b>No distant metastasis</b>		<b>M0</b>	

Scalene (ipsi./ contralateral)	Supraclavicular	Contralateral		Ipsilateral		<b>LYMPH NODE (N)</b>	Stage
		Hilar	Mediastinal	Mediastinal	Hilar		
●	●	●	●			<b>N3</b>	Stage III B
-	-	-	-	●	●	<b>N2</b>	Stage III A
-	-	-	-		●	<b>N1</b>	Stage II B
-	-	-	-	-	-	<b>N0</b>	I A1 I A2 I A3 I B II A II B

**Explanation of lymph node staging:**

- For any N category, one or more of the groups marked by ● must be involved and the involvement of all groups marked by □ should be absent.
- The presence or absence of involvement in groups marked by □ does not alter N staging in the corresponding category.

	<b>PRIMARY TUMOR (T)</b>						
	T1a	T1b	T1c	T2a	T2b	T3	T4
<b>1- Size (greatest dimension)</b>	≤ 1 cm	> 1 cm ≤ 2 cm	> 2 cm ≤ 3 cm	> 3 cm ≤ 4 cm	> 4 cm ≤ 5 cm	> 5 cm ≤ 7 cm	> 7 cm
<b>2- Criteria of Extent</b>				or Any size ≤ 5 cm in the presence of 1 or more of the criteria of extent *		or Any size ≤ 7 cm in the presence of 1 or more of the criteria of extent	or Any size if 1 or more of the criteria of extent are present
<b>Endo-bronchial Location</b>	No extension proximal to the lobar bronchus **			Main bronchus (regardless of the distance to the carina)*** vs. Atelectasis or obstructive pneumonitis extending to the hilum (entire or part of the lung)			Carina or Trachea
<b>Local Invasion</b>	None; the tumor is surrounded by lung or visceral pleura			Visceral pleura		Chest wall (Including superior sulcus), phrenic nerve, parietal pleura and/or parietal pericardium	Diaphragm, Mediastinum, heart, great vessels, recurrent laryngeal nerve, esophagus and/or vertebral body
<b>Separate Tumor Nodule(s)</b>	Absent			Absent		Present in the same lobe of the primary tumor	Present in a different ipsilateral lobe

**Stage I A1 (T1 (mi) N0 M0)**  
T1 (mi): Minimally invasive adenocarcinoma (solitary adenocarcinoma, ≤ 3 cm with a lepidic growth and ≤ 5 mm invasion in any focus)

**Stage 0 (Tis N0 M0)**  
Tis: Carcinoma in situ

**Occult Carcinoma (Tx N0 M0)**  
Tx: Tumor is proven histopathologically (+ Cytology) but not detected by imaging or bronchoscopy)

# TNM

## A Classification for Clinical Cancer Stage Parameters

- most widely used cancer staging system
- **T** refers to the size and extent of the main tumor
- **N** refers to the the number / location of nearby lymph nodes that have cancer infiltration
- **M** refers to whether the cancer has metastasized
- not used for leukemias / lymphomas
  - Binet and Rai in CLL
  - proportion of blasts in bone marrow or blood in leukemias
  - Lugano classification in lymphomas
- other disease specific staging systems may (co-) exist
  - e.g. a stage II lung cancer is determined by size & nodal involvement

TNM has been "ontologized" into NCIt

NCIT:C48698	Cancer TNM Finding Category	0
NCIT:C133398	Postneoadjuvant Therapy Pathologic TNM Finding	1
NCIT:C143081	Posttherapy Clinical TNM Finding	1
NCIT:C48739	Pathologic TNM Finding	1
NCIT:C48886	Pathologic Distant Metastasis TNM Finding	2
NCIT:C48740	pM0 Stage Finding	3
NCIT:C48741	pM1 Stage Finding	3
NCIT:C48742	pM1a Stage Finding	4
NCIT:C48743	pM1b Stage Finding	4
NCIT:C48744	pM1c Stage Finding	4
NCIT:C48887	Pathologic Regional Lymph Nodes TNM Finding	2
NCIT:C48745	pN0 Stage Finding	3
NCIT:C48746	pN1 Stage Finding	3
NCIT:C48747	pN1a Stage Finding	4
NCIT:C48748	pN1b Stage Finding	4
NCIT:C48749	pN1c Stage Finding	4
NCIT:C48750	pN2 Stage Finding	3
NCIT:C48751	pN2a Stage Finding	4
NCIT:C48752	pN2b Stage Finding	4
NCIT:C48753	pN2c Stage Finding	4
NCIT:C48754	pN3 Stage Finding	3
NCIT:C48755	pN3a Stage Finding	4
NCIT:C48756	pN3b Stage Finding	4
NCIT:C48757	pN3c Stage Finding	4
NCIT:C48888	Pathologic Primary Tumor TNM Finding	2
NCIT:C48758	pT0 Stage Finding	3
NCIT:C48759	pT1 Stage Finding	3
NCIT:C48760	pT1a Stage Finding	4
NCIT:C48761	pT1b Stage Finding	4
NCIT:C48763	pT1c Stage Finding	4
NCIT:C48764	pT2 Stage Finding	3
NCIT:C48765	pT2a Stage Finding	4
NCIT:C48766	pT2b Stage Finding	4
NCIT:C48767	pT2c Stage Finding	4
NCIT:C48768	pT3 Stage Finding	3
NCIT:C48769	pT3a Stage Finding	4
NCIT:C48770	pT3b Stage Finding	4
NCIT:C48771	pT3c Stage Finding	4
NCIT:C48772	pT4 Stage Finding	3
NCIT:C48773	pT4a Stage Finding	4
NCIT:C48774	pT4b Stage Finding	4
NCIT:C48775	pT4c Stage Finding	4
NCIT:C48776	pT4d Stage Finding	4
NCIT:C48879	Generic TNM Finding	1
NCIT:C48777	Cancer TNM Vessel Invasion Finding Category	2
NCIT:C147091	Lymphovascular Invasion 0	3
NCIT:C147092	Lymphovascular Invasion 1	3
NCIT:C147093	Lymphovascular Invasion 9	3
NCIT:C147094	Lymphovascular Invasion 2	3
NCIT:C147095	Lymphovascular Invasion 3	3
NCIT:C147096	Lymphovascular Invasion 4	3
NCIT:C48883	Generic Distant Metastasis TNM Finding	2
NCIT:C48699	M0 Stage Finding	3
NCIT:C95956	cM0 (+) Stage Finding	4
NCIT:C48700	M1 Stage Finding	3
NCIT:C48701	M1a Stage Finding	4
NCIT:C48702	M1b Stage Finding	4
NCIT:C48703	M1c Stage Finding	4

NCIT:C48704	MX Stage Finding	0
NCIT:C48884	Generic Regional Lymph Nodes TNM Finding	1
NCIT:C48705	N0 Stage Finding	1
NCIT:C95921	N0 (-) Stage Finding	1
NCIT:C95922	N0 (+) Stage Finding	2
NCIT:C95923	N0 (mol-) Stage Finding	3
NCIT:C95925	N0 (mol+) Stage Finding	3
NCIT:C48706	N1 Stage Finding	4
NCIT:C48707	N1a Stage Finding	4
NCIT:C48708	N1b Stage Finding	4
NCIT:C95929	N1bl Stage Finding	2
NCIT:C95935	N1blI Stage Finding	3
NCIT:C95936	N1blII Stage Finding	3
NCIT:C95937	N1blIV Stage Finding	4
NCIT:C48709	N1c Stage Finding	4
NCIT:C95955	N1mi Stage Finding	4
NCIT:C48714	N3 Stage Finding	3
NCIT:C48715	N3a Stage Finding	4
NCIT:C48716	N3b Stage Finding	4
NCIT:C48717	N3c Stage Finding	4
NCIT:C48718	NX Stage Finding	3
NCIT:C48786	N2 Stage Finding	4
NCIT:C48711	N2a Stage Finding	4
NCIT:C48712	N2b Stage Finding	4
NCIT:C48713	N2c Stage Finding	2
NCIT:C96026	N4 Stage Finding	3
NCIT:C48885	Generic Primary Tumor TNM Finding	3
NCIT:C106299	Any T	4
NCIT:C132010	T5 Stage Finding	4
NCIT:C48719	T0 Stage Finding	4
NCIT:C48720	T1 Stage Finding	3
NCIT:C48721	T1a Stage Finding	4
NCIT:C48722	T1b Stage Finding	4
NCIT:C48723	T1c Stage Finding	4
NCIT:C95805	T1mi Stage Finding	3
NCIT:C48724	T2 Stage Finding	4
NCIT:C148411	T2d Stage Finding	4
NCIT:C48725	T2a Stage Finding	4
NCIT:C48726	T2b Stage Finding	3
NCIT:C48727	T2c Stage Finding	4
NCIT:C48728	T3 Stage Finding	4
NCIT:C148412	T3d Stage Finding	4
NCIT:C48729	T3a Stage Finding	4
NCIT:C48730	T3b Stage Finding	1
NCIT:C48731	T3c Stage Finding	2
NCIT:C48732	T4 Stage Finding	3
NCIT:C48733	T4a Stage Finding	3
NCIT:C48734	T4b Stage Finding	3
NCIT:C48735	T4c Stage Finding	3
NCIT:C48736	T4d Stage Finding	3
NCIT:C48737	TX Stage Finding	3
NCIT:C48738	Tis Stage Finding	2
NCIT:C96025	Ta Stage Finding	3
NCIT:C48880	Recurrent Cancer TNM Finding	4
NCIT:C48881	Clinical TNM Finding	3
NCIT:C161009	Clinical Primary Tumor TNM Finding	4
NCIT:C162609	Clinical Regional Lymph Nodes TNM Finding	4
NCIT:C162610	Clinical Distant Metastasis TNM Finding	4
NCIT:C48882	Autopsy TNM Finding	

# Tasks

## Survival analyses | Cancer classifications | Staging

- Familiarize yourself with the different concepts behind different disease classification systems - what are their use, advantages, problems? E.g. ICD-10, ICD-O, NCIIt
  - you can use Progenetix to explore e.g. ontology mapping
- Learn to "read" Kaplan-Meier plots (preparation for explorative analyses later this week).
- Achieve a principal understanding of TNM codes & write some "translations"
  - ➔ T1N1M0: small tumor with regional lymph node involvement and no detected distant metastases