A Hidden Treasure of Data Science

Adverse Event (AE)

Dung-Tsa Chen, PhD
Zachary Thompson, PhD
Junmin Whiting, PhD
Alyssa Obermayer, MS
Melanie Buhlmann, MS
Jiannong Li, PhD
Ram Thapa, PhD
Timothy Shaw, PhD

*Funding support by NCI (1R21CA286417), Florida Cancer Innovation Fund (MOAAX), and Moffitt Biostatistics and Bioinformatics Department fund

AE Translational Research Team

Data Science Experts

Biostatistics & Bioinformatics

Dr. Dung-Tsa Chen (PI)

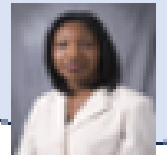
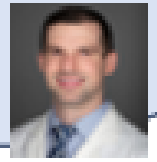
Dr. Timothy Shaw


Clinical Experts

Thoracic Oncology

Dr. Jhanelle Gray (Co-PI)

Dr. Andreas Saltos

Immunology Expert

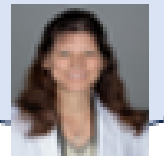
Immunology

Dr. Amer Beg

Pathology Expert

Pathology

Dr. Theresa Boyle

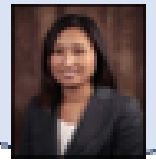


AE Predictive Biomarkers

Clinical Expert

Breast Oncology

Dr. Hyo Han



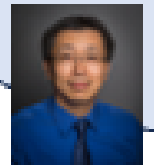
Clinical Experts

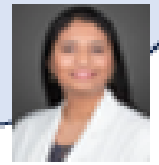
GI Oncology

Dr. Richard Kim

Dr. Dae Won Kim

Dr. Rutika Mehta

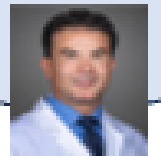




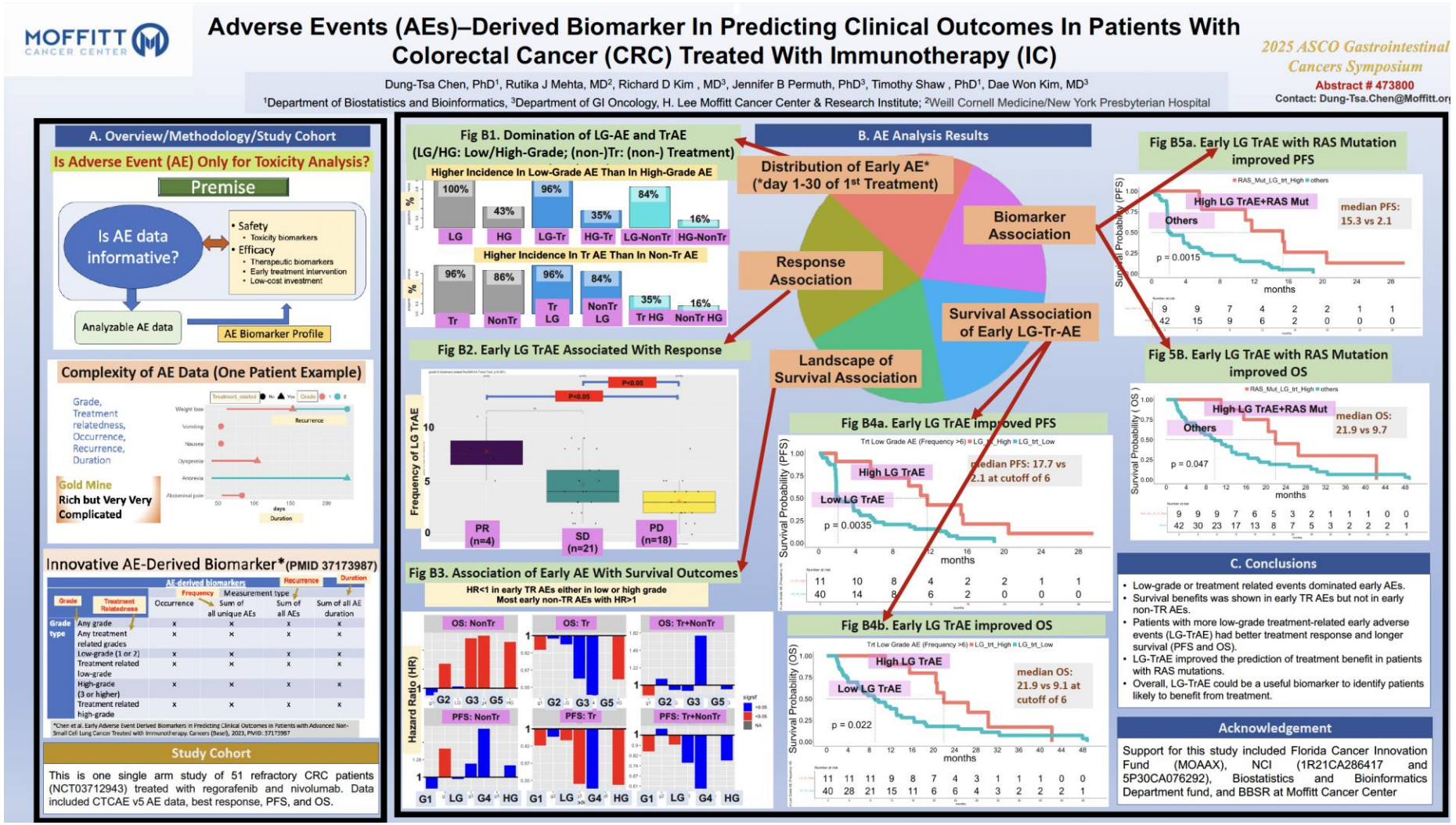
Clinical Expert

Cutaneous Oncology

Dr. Ahmad Tarhini



Evidences of Hidden Treasure (1)



Evidences of Hidden Treasure (2)

Preview Of Early-Onset Adverse Event (AE) Landscape And Association With Clinical Outcomes In Advanced Non-Small Cell Lung Cancer (NSCLC)

Dung-Tsa Chen, Andreas N. Saltos, Zachary J. Thompson, Timothy Shaw, Jhanelle E. Gray
Moffitt Cancer Center

MOFFITT CANCER CENTER

AACR 2025 # 3365

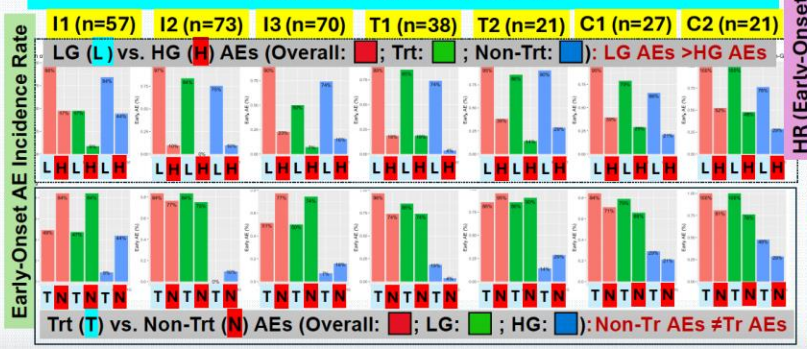
Introduction

- Early-onset AE incidence and the association with clinical outcomes not well defined in advanced NSCLC.
- Most published trial studies are limited in using worst grade approach for AE report.
- A need to delineate Early-onset AE landscape and explore Early-onset AE potential role as a prognostic factor.

Study Design

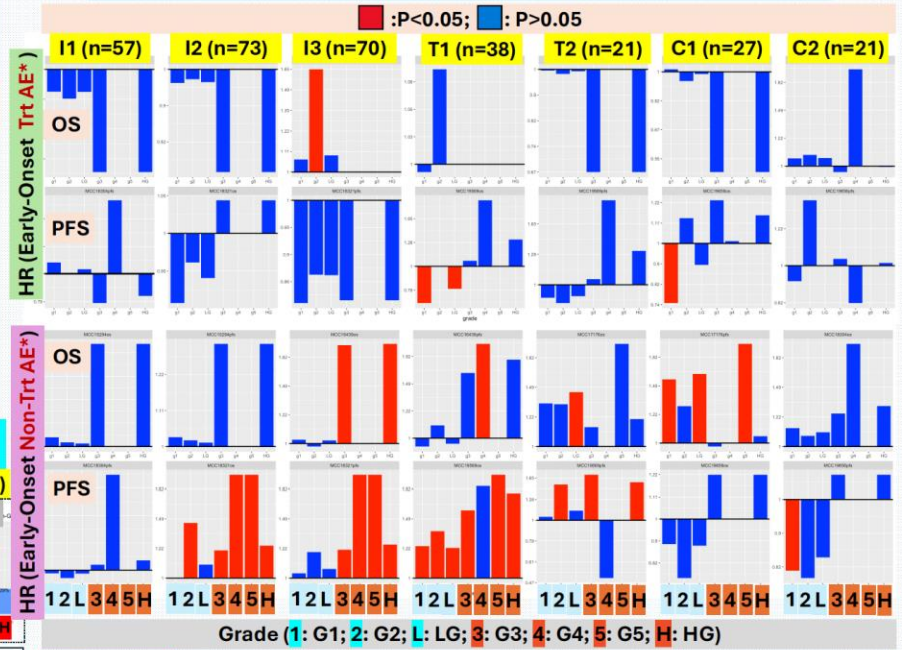
- All AEs were coded according to National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE).
- **Low-Grade (LG) AE**: Grade 1-2; **High Grade (HG) AE**: Grade 3-5.
- **Treatment related (Trt) AE**: Attribution of 'definitely', 'probable', or 'possible'.
- **Non-Treatment (Non-Trt) AE**: Attribution of 'unrelated' or 'unlikely'.
- **Early-onset AE**: AE occurrence since first treatment until day 30.
- Early-onset AE incidence rate: Percentage of patients with AE experiences.
- Early-onset Trt and Non-Trt AEs (frequency) were derived by AE-derived biomarkers*.
- **Study cohorts**: 3 immunotherapy cohorts :(I1 Druva+mRNA; I2: GM.CD40L_CCL21; I3: PBF509_PDR001); 2 targeted therapy cohorts (T1: Cetuximab_Radiotherapy; T2: Afatinib_Dasatinib), 2 chemotherapy cohorts (C1: AZD1775_Carbo_Pacli; C2: Ceritinib+Docetaxel) in advanced NSCLC.

Early-Onset AE Landscape



Association of Early-Onset AEs With Survival Outcomes

Trt: HR<1 in most LG AEs; Non-Trt: HR>1 in most LG and HG AEs



Conclusion. This is the first study to portray early-onset AE outlook in advanced NSCLC. Further analysis of survival association in early-onset AEs identified low-grade treatment related early-onset AEs correlated with better PFS or OS.

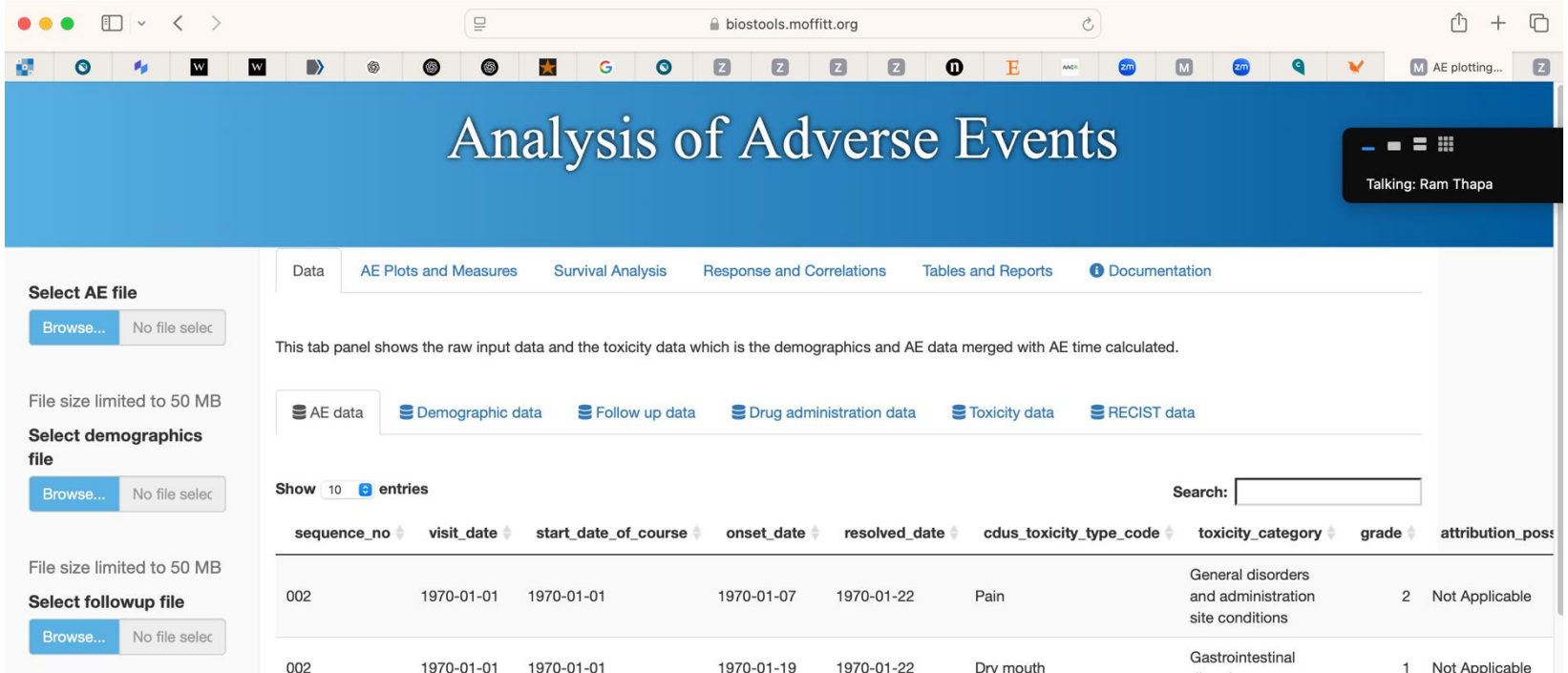
Funding Support: NCI (1R21CA286417), Florida Cancer Innovation Fund (MOAAX), and Moffitt Biostatistics and Bioinformatics Department fund.

*: Chen et al. PMID: **37173987.**

Evidences of Hidden Treasure (3)

Zach's Masterpiece

- R package: AdverseEvents (https://cran.r-project.org/web/packages/AdverseEvents/index.html)
- Biostools https://biostools.moffitt.org/DungTsaChen/AdverseEvents/
- A manuscript in developmental stage



Just Beginning of AE Exploration

Question

Is Adverse Event Really A Bad Thing?

Could Be Some AEs Informative of Clinical Outcomes?

What is Adverse Event (AE)?

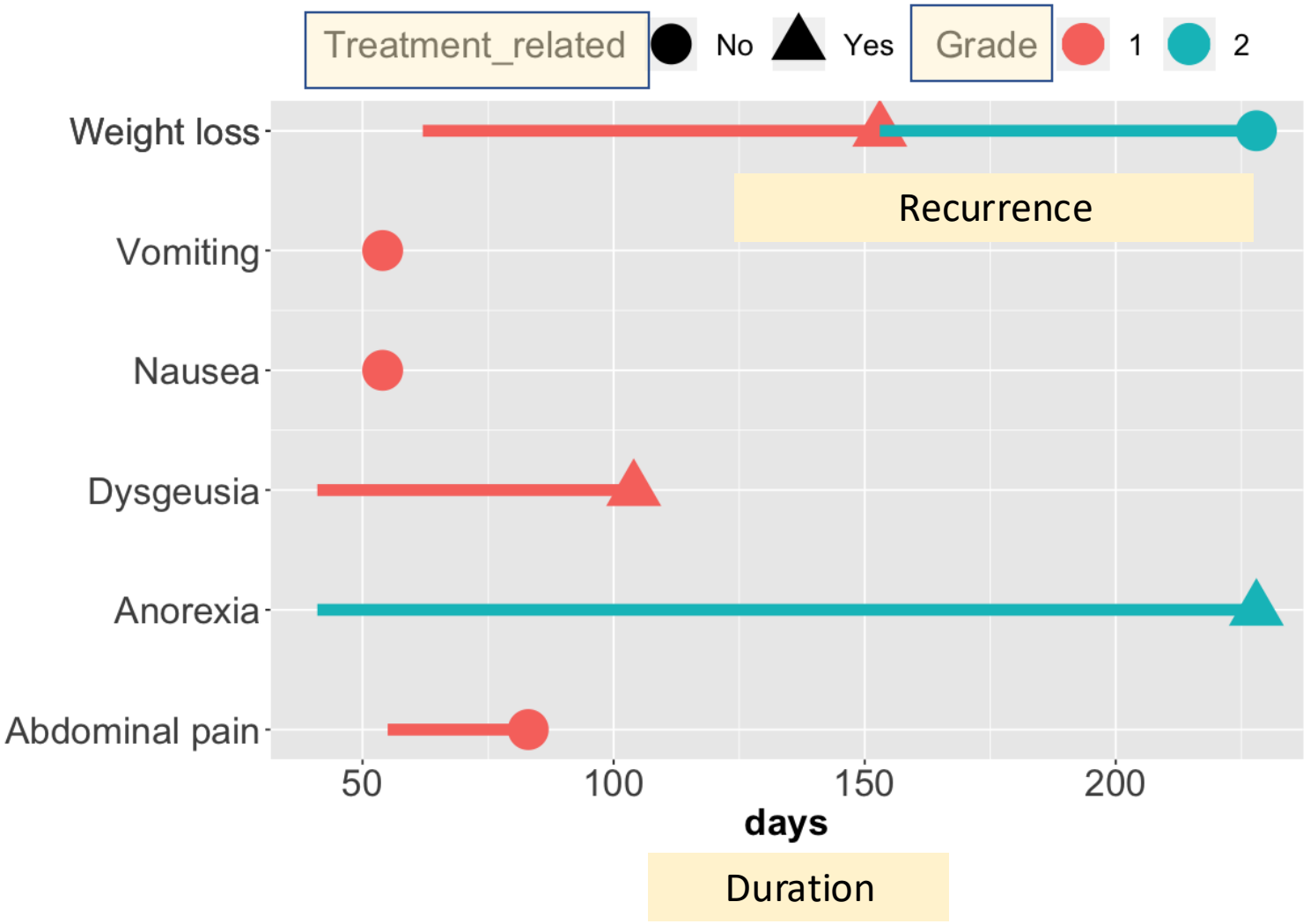
- AE definition (by NCI): **any unfavorable symptom, sign, or disease** (including an abnormal laboratory finding) temporally associated with the use of a medical treatment or procedure that **may or may NOT be related to the medical treatment**.
- A required component to assess patient safety in cancer clinical trial.

Illustrative AE Data of One Lung Cancer Patient

Grade,
Treatment relatedness,
Occurrence,
Recurrence,
Duration

Gold Mine

Rich but Very Very Complicated



AE: A Hidden Treasure

Clinical Innovation

- Toxicity biomarkers
- Therapeutic biomarkers
- Early treatment intervention
- Low-cost but significant improvements in cancer care especially for routine non-research cares

Methodological Innovation

- Paradigm shift of AE data usage from descriptive summary into modern informative AE biomarkers to fulfill precision medicine
- Diverse coverage of AE contents to unlock their potential for clinical application
- Predictive value of early AE biomarkers
- Global discovery AE analysis

Premise

Is AE data informative?

- Safety
  - Toxicity biomarkers
- Efficacy
  - Therapeutic biomarkers
  - Early treatment intervention
  - Low-cost
  - …

Analyzable AE data

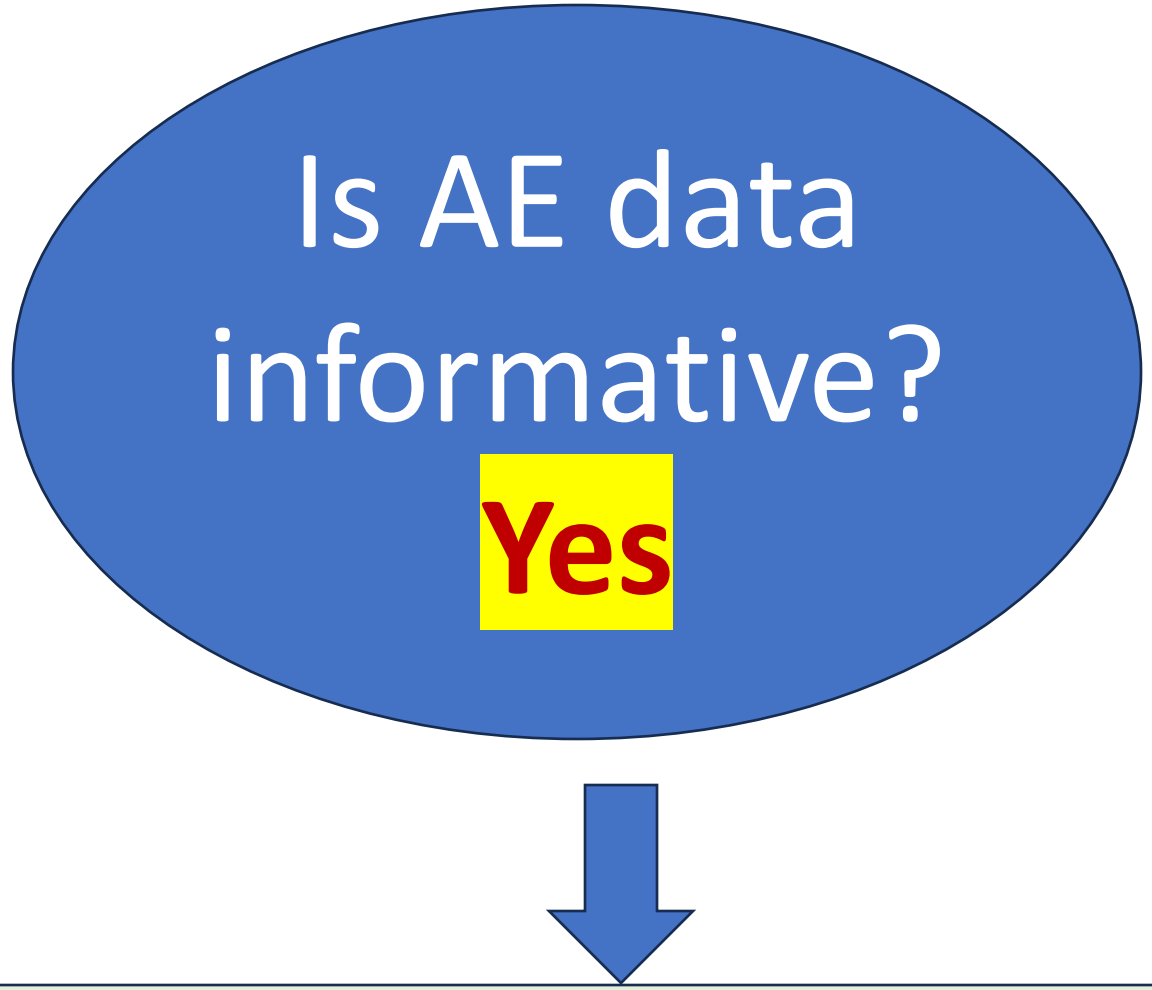
AE Biomarker Profile

Potential Clinical Utility

Association with clinical outcomes in various studies:

- High neutrophil-to-lymphocyte ratio (NLR) correlated with worse survival in multiple immune checkpoint inhibitors (ICI) treated malignancies.
- Hypertension associated with improved survival outcomes
  - Advanced NSCLC patients treated with bevacizumab in combination with carboplatin and paclitaxel (immunotherapy+ chemotherapy)
  - Patients with metastatic renal cell carcinoma treated with sunitinib (targeted therapy).
- Immune-related adverse event (irAE) associated with improved overall survival
  - Melanoma patients treated with nivolumab.
  - NSCLC treated with ICI

- Limitation: restrict to a small AE subset, such as irAE, or not take into account of AE frequency, duration, or grade


Is AE data informative?
Yes
How to make ”analyzable” AE data

How AE Is Measured?

- CTCAE (by NCI): Common Terminology Criteria for Adverse Events (CTCAE) to standardize toxicity.
- The CTCAE version 5 lists 26 toxicity categories to cover a total of 837 AE terms.
- Each AE term includes date of treatment, onset date, resolved date, grade (severity), treatment attribution (treatment relatedness status), and so on.

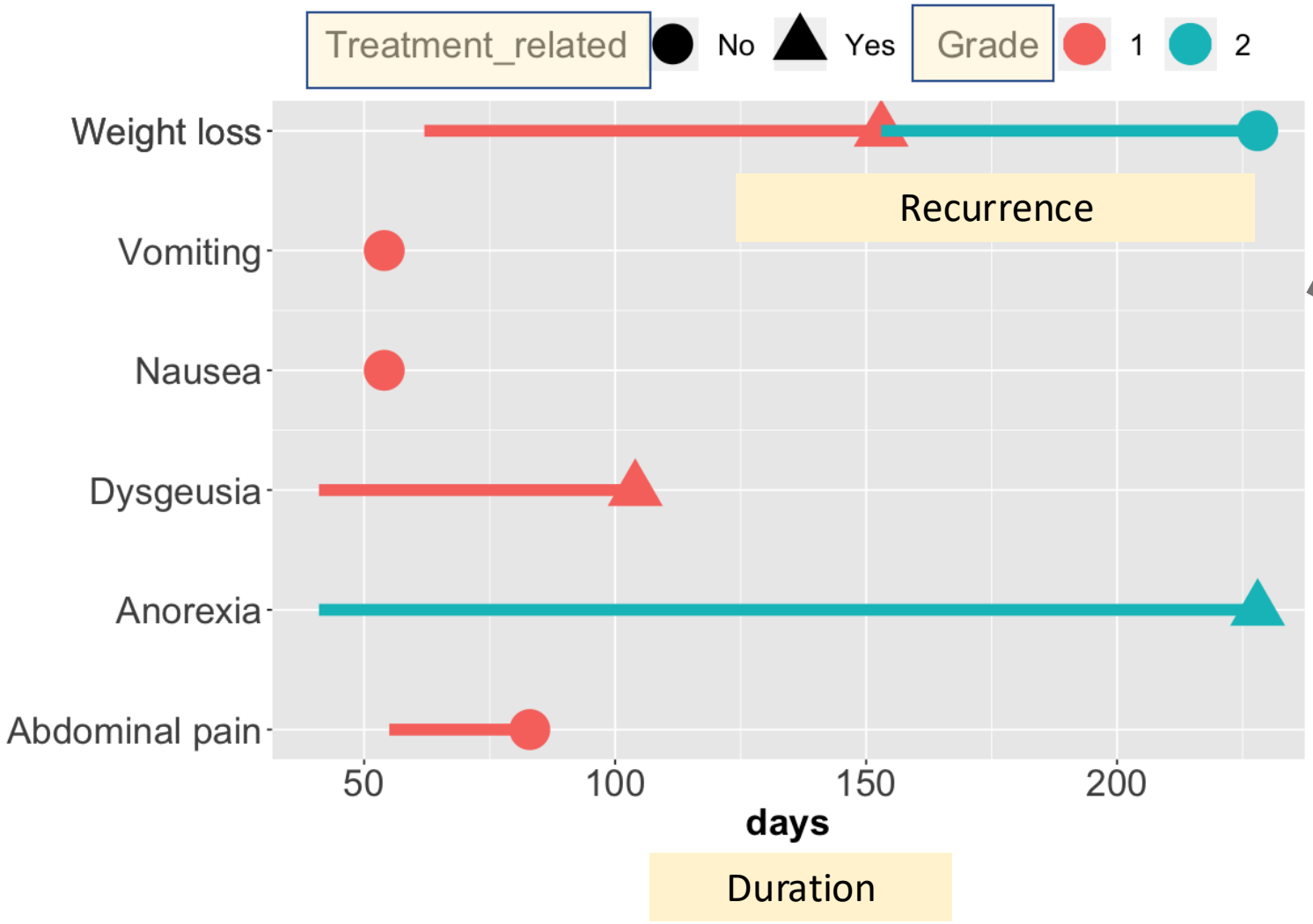
Challenges of AE Data Analysis

- The CTCAE version 5 lists 26 toxicity categories to cover a total of 837 AE terms.

Will the ~800 AE terms be just a n x 837matrix?

	AE_1	AE_2	…	AE_837
Pt_1				
…		??		
Pt_n				

Illustrative AE Data of One Patient (Cont.)



??

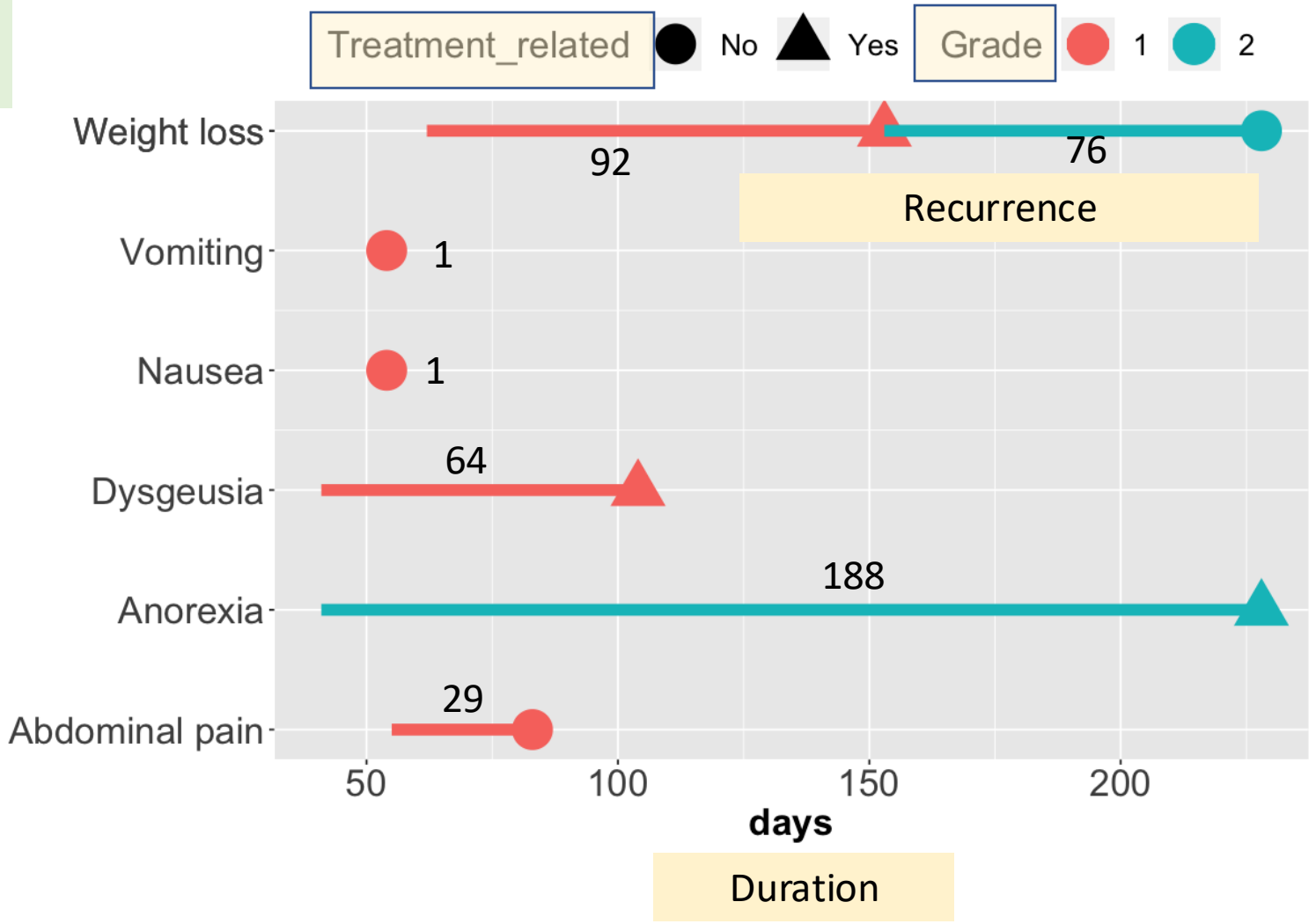
A n x 837 matrix

	AE_1	AE_2	…	AE_837
Pt_1				
…				
Pt_n				

Illustrative AE Data of One Patient (Cont.)

Many ways to dictate AE

- Any AE: **Yes**
- Low grade AE (grade 1-2): **Yes**
- High grade AE (grade 3-5): **No**
- Any treatment related grade AE: **Yes**
- Treatment related low grade AE: **Yes**
- Treatment related high grade AE: **No**
- Frequency of AE: **6**
- Duration of all AEs: **451 days**
- ….



Classical AE Report

Descriptive Statistics unfitted to Precision Medicine Era

One TIL Trial from Nature Medicine

Tumor-infiltrating lymphocyte treatment for anti-PD-1 resistant metastatic lung cancer: a phase I trial

Table 1:

Treatment-emergent adverse events reported with cyclophosphamide, fludarabine, TIL, or interleukin-2.

	Grade					
NCI CTCAE Preferred Term	1	2	3	4	5	Any
Lymphocyte count decreased	0 (0%)	0 (0%)	0 0%)	16 (100%)	0 (0%)	16 (100%)
White blood cell count decreased	0 (0%)	0 (0%)	1 (6%)	15 (94%)	0 (0%)	16 (100%)
Anemia	1 (6%)	2 13%)	13 (81%)	0 (0%)	0 0%)	16 (100%)
Platelet count decreased	0 (0%)	1 (6%)	3 19%)	11 (69%)	0 (0%)	15 (94%)
Neutrophil count decreased	0 (0%)	2 13%)	1 6%)	11 (69%)	0 0%)	14 (88%)
Hypoalbuminemia	0 (0%)	13 (81%)	1 (6%)	0 (0%)	0 (0%)	14 (88%)
Nausea	7 (44%)	7 44%)	0 0%)	0 (0%)	0 0%)	14 (88%)

One Immunotherapy Trial from Lancet

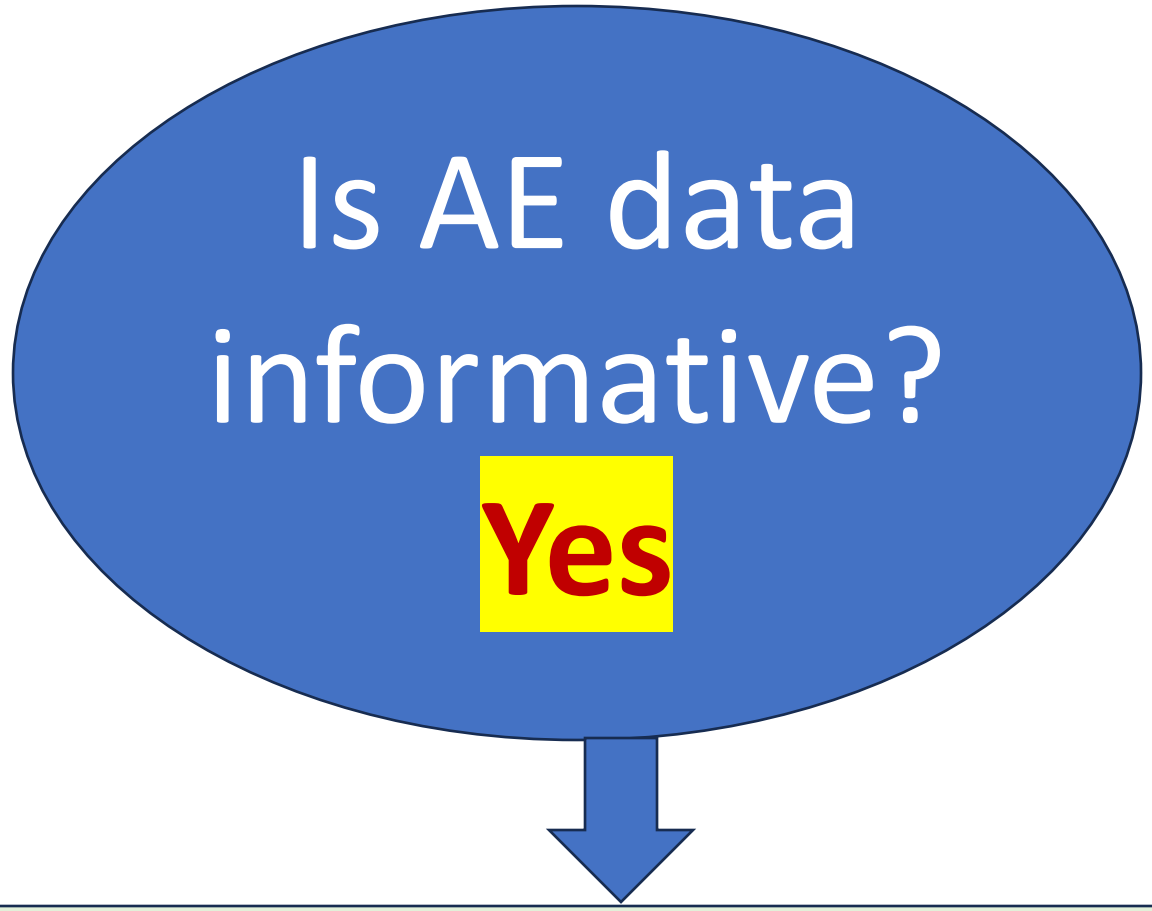
Combined nivolumab and ipilimumab with or without stereotactic body radiation therapy for advanced Merkel cell carcinoma: a randomised, open label, phase 2 trial

Supplementary Table 1. Treatment-related Adverse Events.

	Total Cohort (N=50)			Arr
Adverse Event*	Grade 1-2	Grade 3	Grade 4	Grade 1-2
Any	27 (54%)	13 (26%)	5 (10%)	11 (44%)
Fatigue	31 (62%)	1 (2%)	0	18 (72%)
Pruritus	20 (40%)	0	0	11 (44%)
Diarrhea	20 (40%)	2 (4%)	0	9 (36%)
Rash/Dermatitis	19 (38%)	1 (2%)	0	11 (44%)
Nausea	16 (32%)	0	0	9 (36%)
Elevated pancreatic enzymes	14 (28%)	2 (4%)	4 (8%)	12 (48%)
Cough	12 (24%)	0	0	6 (24%)
Anorexia	11 (22%)	0	0	3 (12%)

Some Existing Statistical Approaches

- ToxT approach: Incorporate the dimension of time into AE assessment to depict toxic effects.
  - Limitation: impractical mean of AE grade and fixed time point by drug cycle.
- Generalized log-rank test: Use mean frequency function to compare recurrent AE events.
  - Limitation: lack of consideration of AE duration
- Q-TWiST approach: Analyze quality adjusted time without symptoms or toxicity by decomposing survival time into three health states: time with toxicity (TOX), time without toxicity, relapse, or progression (TWiST), and time after tumor progression or relapse until death (REL), and then uses a weighted average of the three health states to form the Q-TWiST score.
  - Limitation: dependency of utility weights and inability in individual AE analysis.



How to make ”analyzable” AE data

Efficient & Effective

Our Innovative Analyzable AE Data

- **Integration of different AE parameters**
  - Grade
  - Treatment relatedness
  - Occurrence
  - Recurrence
  - Duration
- **Conversion into a series of analyzable AE metrics (AE biomarkers)**
  - Occurrence
  - Unique frequency
  - Frequency including recurrence
  - Duration

Our Unique Framework of AE Analysis

- **Utilization of AE parameters** to derive AE biomarkers by toxicity severity level (grade), treatment relatedness, AE occurrence, frequency, and duration.
- **Comprehensive analysis** of AE-derived biomarkers from overall AE, toxicity category, down to individual AE.
- **Early AE analysis** at day 30 from initial treatment date for early prediction.
- **Informative analysis components** including survival plot and boxplot for each AE-derived biomarker, summary plots of effect size and p value for all AE-derived biomarkers in an AE, summary tables for significant AEs.

Innovative AE-Derived Biomarker*

AE-derived biomarkers					
Grade	Treatment Relatedness	Measurement type			
		Occurrence	Sum of all unique AEs	Sum of all AEs	Sum of all AE duration
Grade type	Any grade	x	x	x	x
	Any treatment related grades	x	x	x	x
	Low-grade (1 or 2)	x	x	x	x
	Treatment related low-grade	x	x	x	x
	High-grade (3 or higher)	x	x	x	x
	Treatment related high-grade	x	x	x	x

Frequency

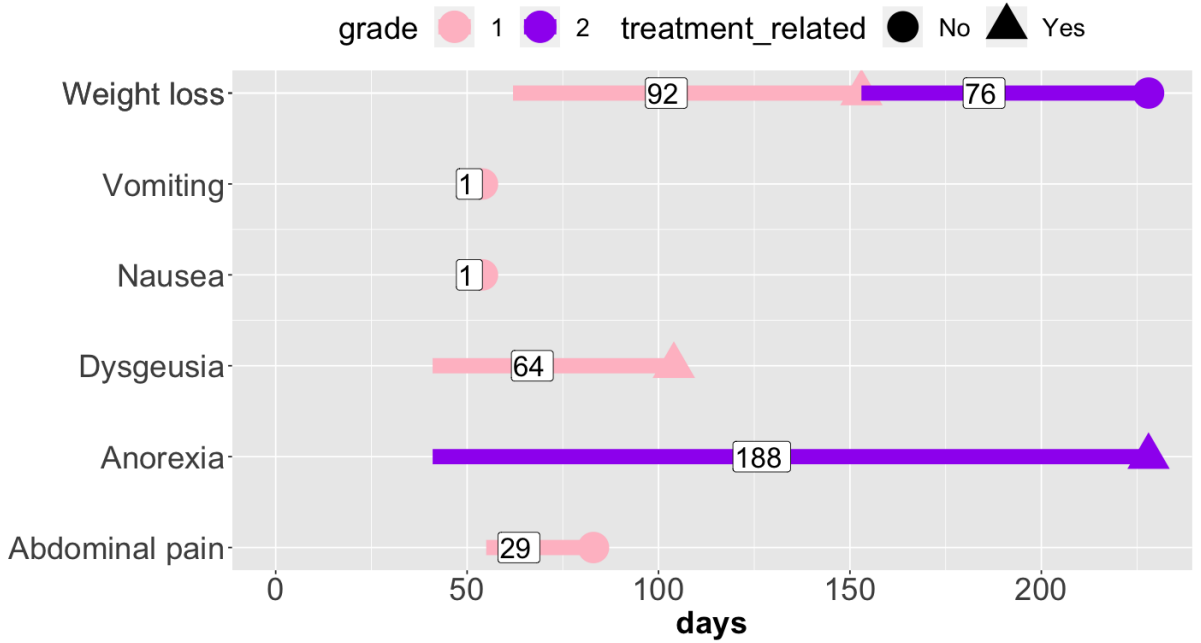
Recurrence

Duration

*Chen et al. Early Adverse Event Derived Biomarkers in Predicting Clinical Outcomes in Patients with Advanced Non-Small Cell Lung Cancer Treated with Immunotherapy. Cancers (Basel), 2023, PMID: 37173987

Example 1: Overall AE

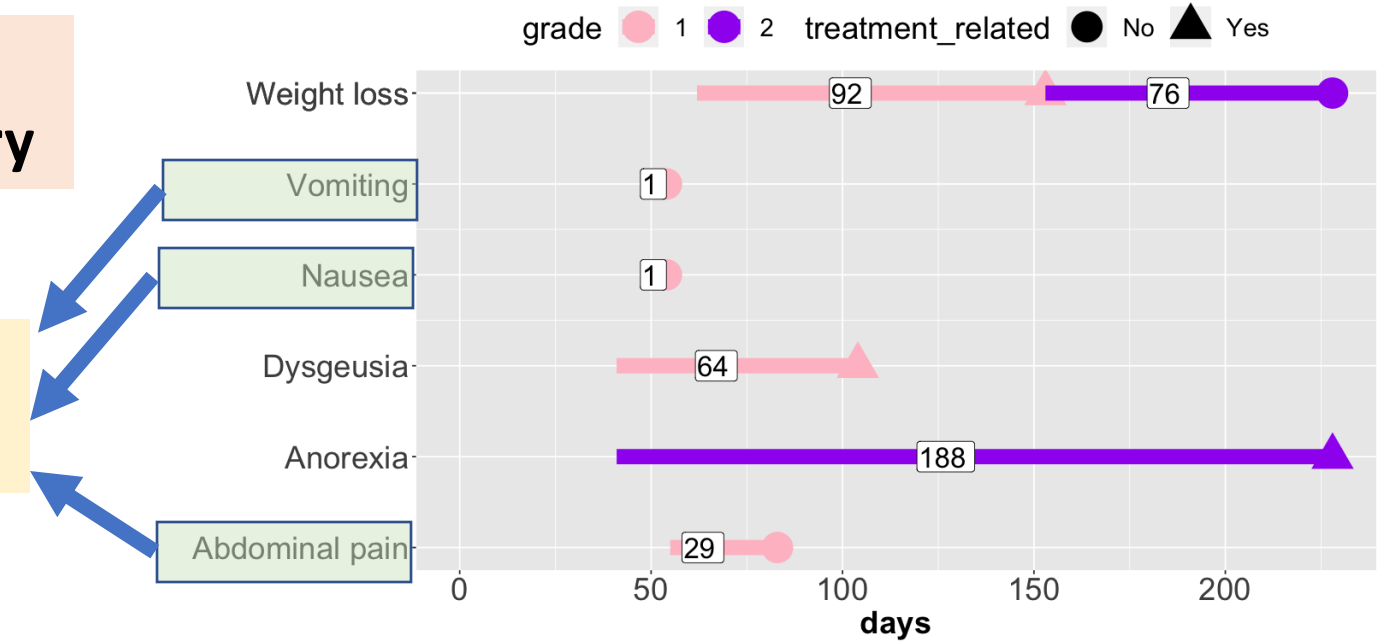
Spareness



AE-derived biomarkers (Patient A for entire AE level)					
		Measurement type			
		Occurrence	Sum of all unique AEs	Sum of all AEs	Sum of all AE duration (days)
Grade/ Treatment relatedness	Any grade	1	6	7	451
	Any treatment related grade	1	3	3	344
	Low-grade (1 or 2)	1	6	7	451
	Treatment related low-grade	1	3	3	344
	High-grade (3 or higher)	0	0	0	0
	Treatment related high-grade	0	0	0	0

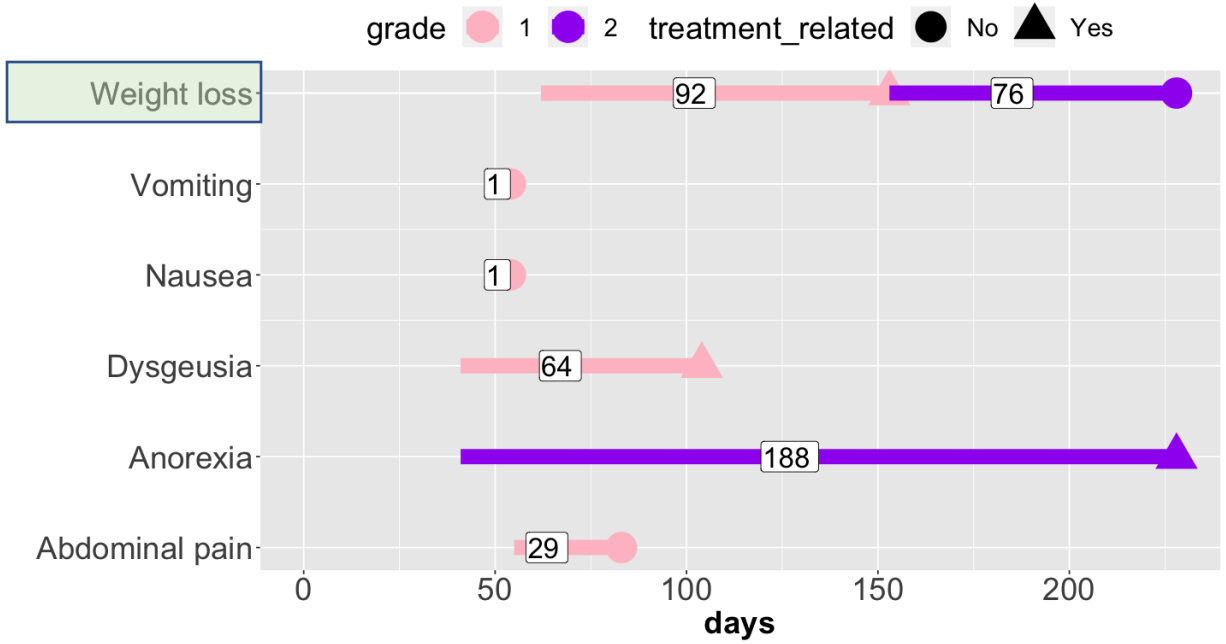
Example 3: Toxicity category

Gastrointestinal disorders



AE-derived biomarkers (Patient A for gastrointestinal disorders)					
		Measurement type			
		Occurrence	Sum of all unique AEs	Sum of all AEs	Sum of all AE duration (days)
Grade/ Treatment relatedness	Any grade	1	3	3	31
	Any treatment related grade	0	0	0	0
	Low-grade (1 or 2)	1	3	3	31
	Treatment related low-grade	0	0	0	0
	High-grade (3 or higher)	0	0	0	0

Example 4: Individual AE



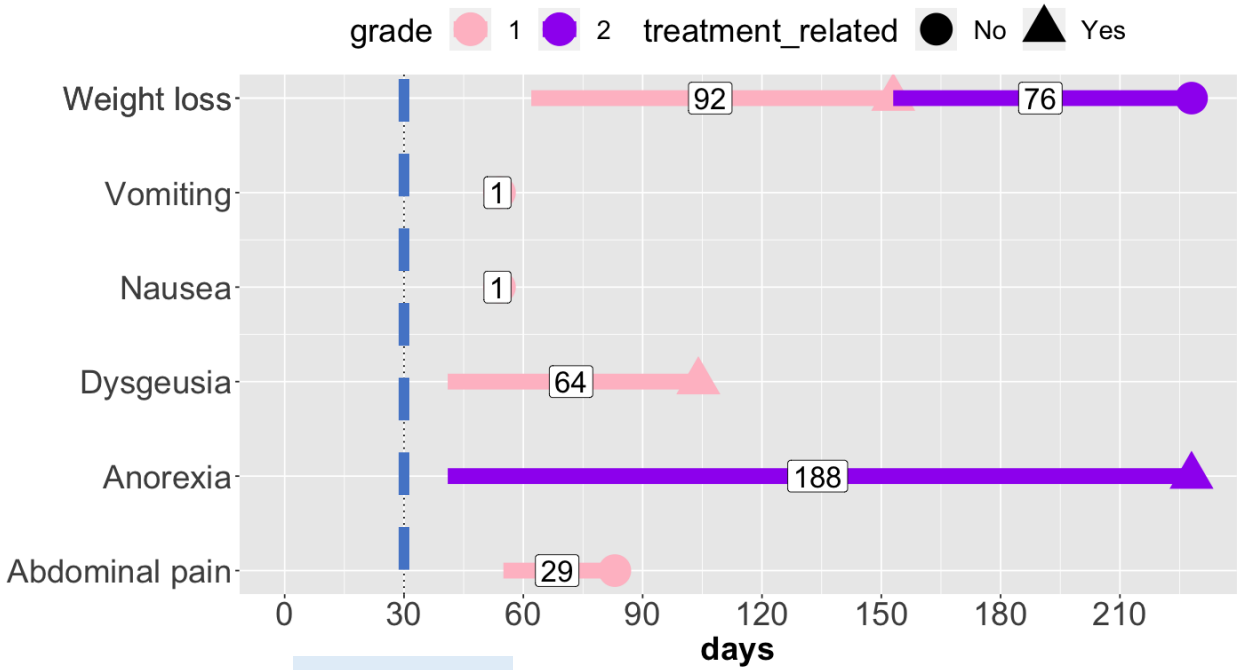
AE-derived biomarkers (Patient A for weight loss)

		Measurement type			
		Occurrence	Sum of all unique AEs	Sum of all AEs	Sum of all AE duration (days)
Grade/ Treatment relatedness	Any grade	1	1	2	168
	Any treatment related grade	1	1	1	92
	Low-grade (1 or 2)	1	1	2	168
	Treatment related low-grade	1	1	1	92
	High-grade (3 or higher)	0	0	0	0
	Treatment related high-grade	0	0	0	0

Early AE Analysis

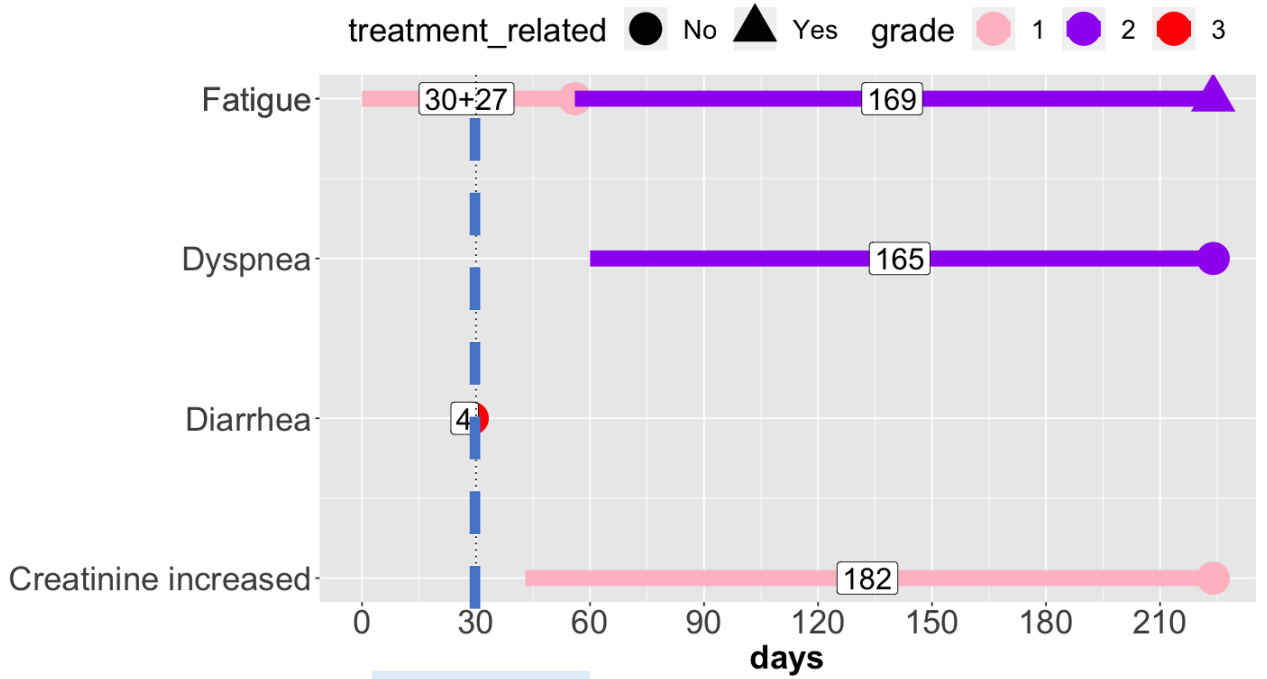
Day 30 as example

No Early AE



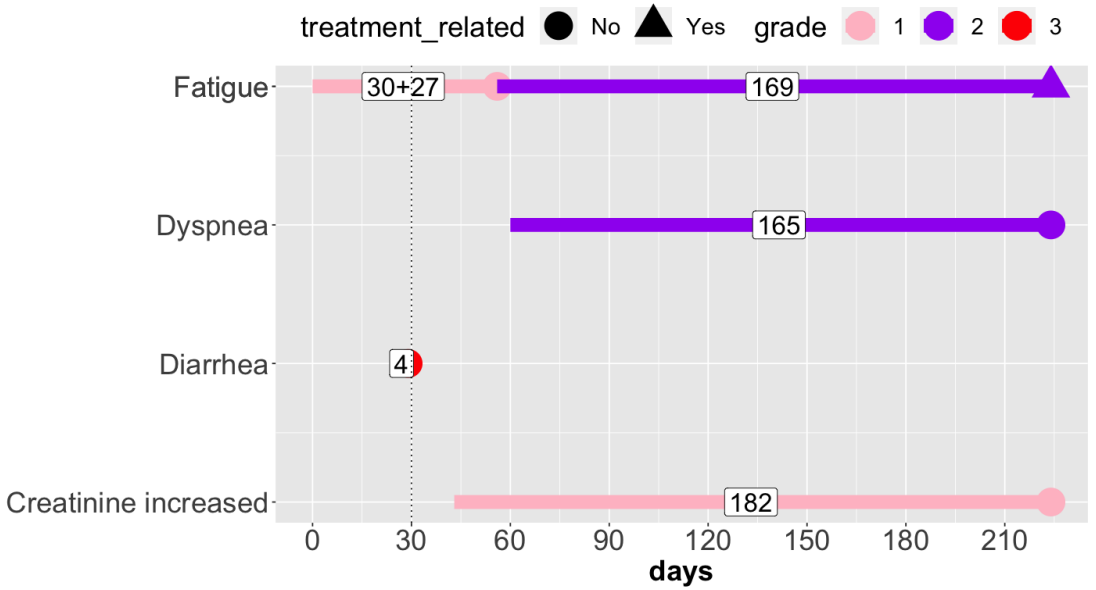
Day 30

Early AEs



Day 30

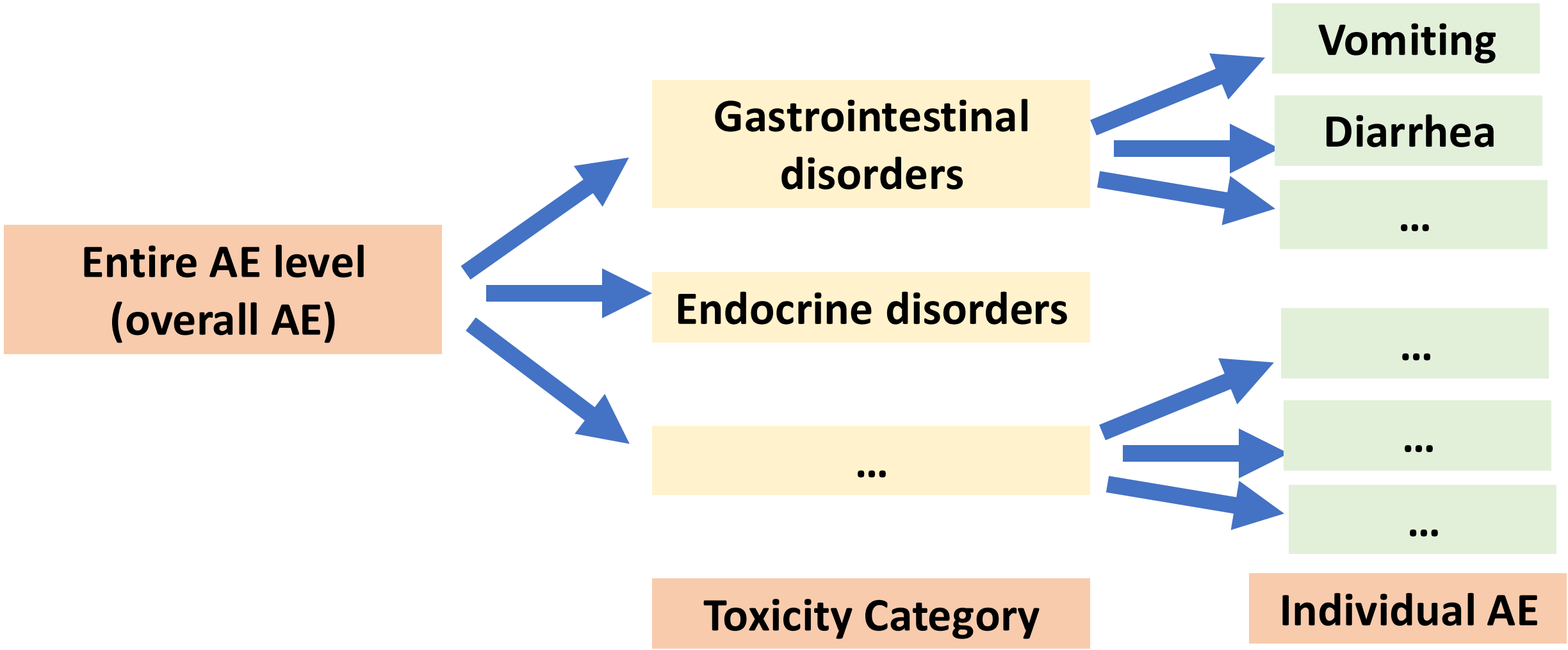
Example: Overall Early AE



Early AE-derived biomarkers (Patient B): Overall AE

		Measurement type			
		Occurrence	Sum of all unique AEs	Sum of all AEs	Sum of all AE duration (days)
Grade/ Treatment relatedness	Any grade	1	2	2	34
	Any treatment related grade	0	0	0	0
	Low-grade (1 or 2)	1	1	1	30
	Treatment related low-grade	0	0	0	0
	High-grade (3 or higher)	1	1	1	4
	Treatment related high-grade	0	0	0	0

Global Discovery AE Analysis



Data Example

- An ongoing phase I/II immunotherapy trial in lung cancer (NCT02638090)
- Population: late-stage non-small cell lung cancer (n=32 patients)
- Treatment: combination of HDAC inhibitor and immunotherapy
- Survival outcomes: overall survival (OS) and progression free survival (PFS)

Survival Association of AE-Derived Biomarker

AE-derived biomarkers					
		Measurement type			
		Occurrence	Sum of all unique AEs	Sum of all AEs	Sum of all AE duration
Grade type	Any grade	x	x	x	x
	Any treatment related grades	x	x	x	x
	Low-grade (1 or 2)	x	x	x	x
	Treatment related low-grade	x	x	x	x
	High-grade (3 or higher)	x	x	x	x
	Treatment related high-grade	x	x	x	x

*Chen et al. Early Adverse Event Derived Biomarkers in Predicting Clinical Outcomes in Patients with Advanced Non-Small Cell Lung Cancer Treated with Immunotherapy. Cancers (Basel), 2023, PMID: 37173987

Simple Survival Analysis

Univariate Cox model

for overall survival (OS) and progress-free survival (PFS)

$$h_i(t) = h_0(t)\exp(\boldsymbol{\alpha} * \boldsymbol{Early\ AE\ biomarker\ metric})$$

Summary of AE-Derived Biomarker Analysis

Treatment Related High Grade AEs

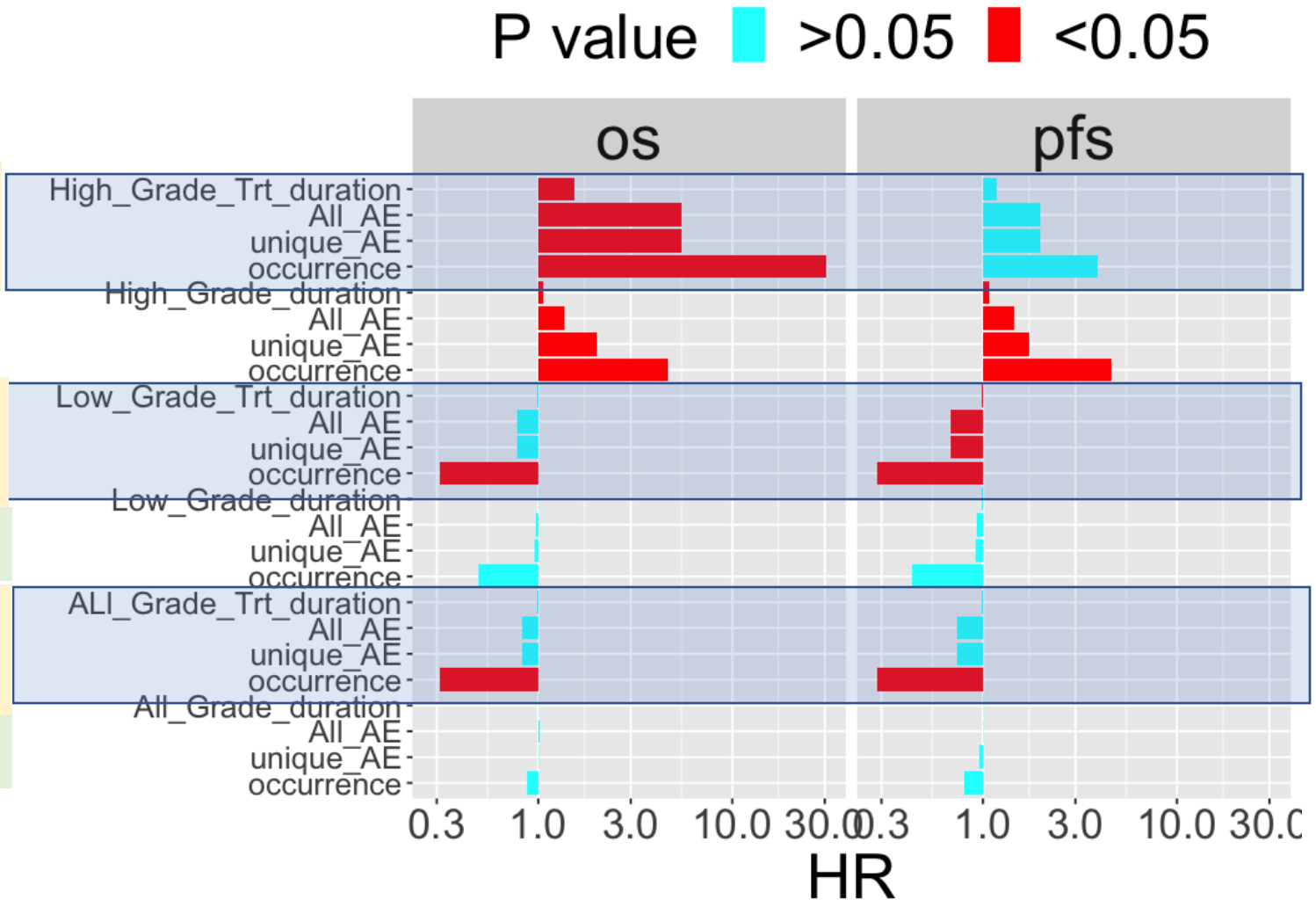
High Grade AEs

Treatment Related Low Grade AEs

Low Grade AEs

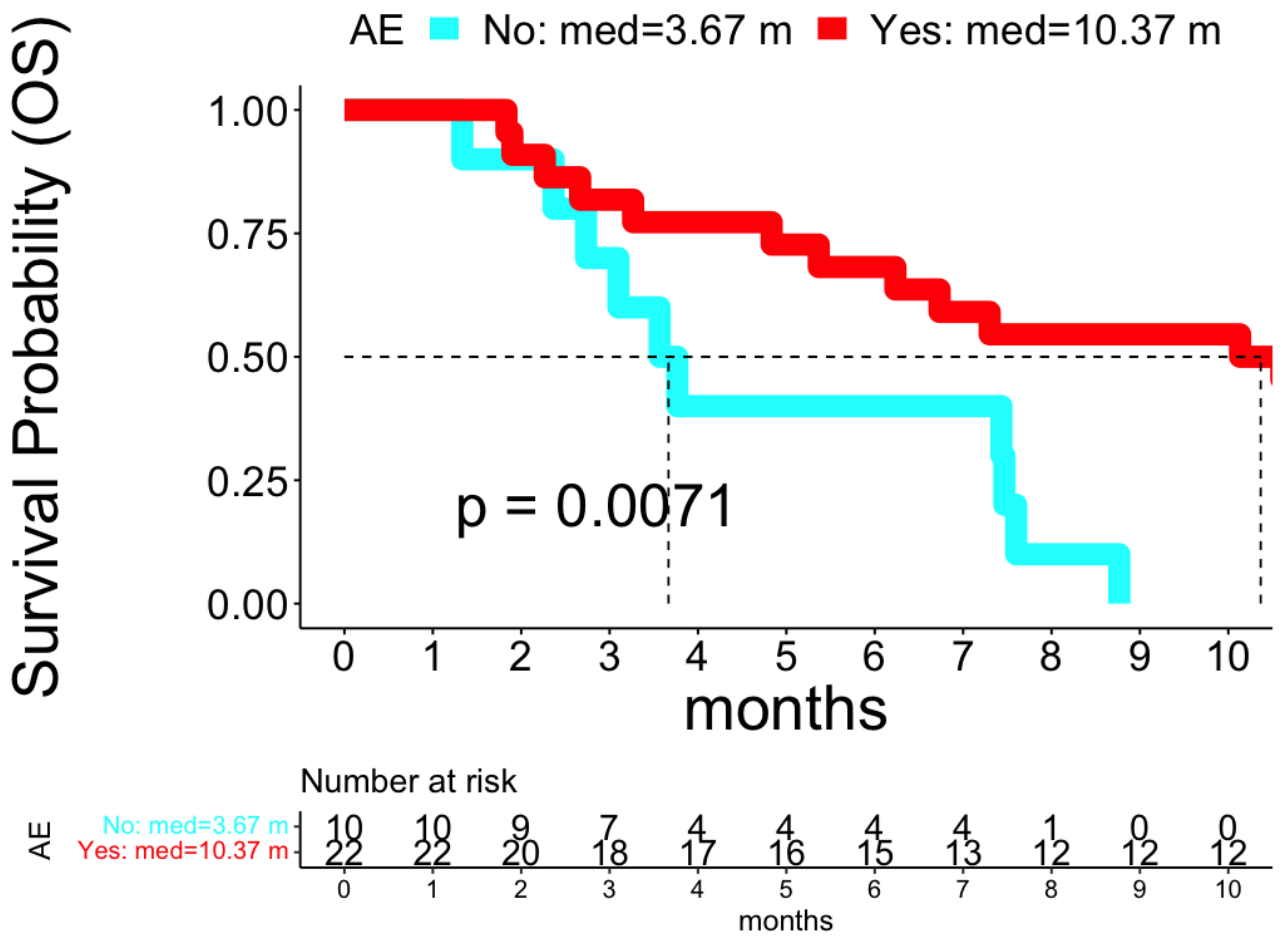
Treatment Related Any Grade AEs

Any Grade AEs



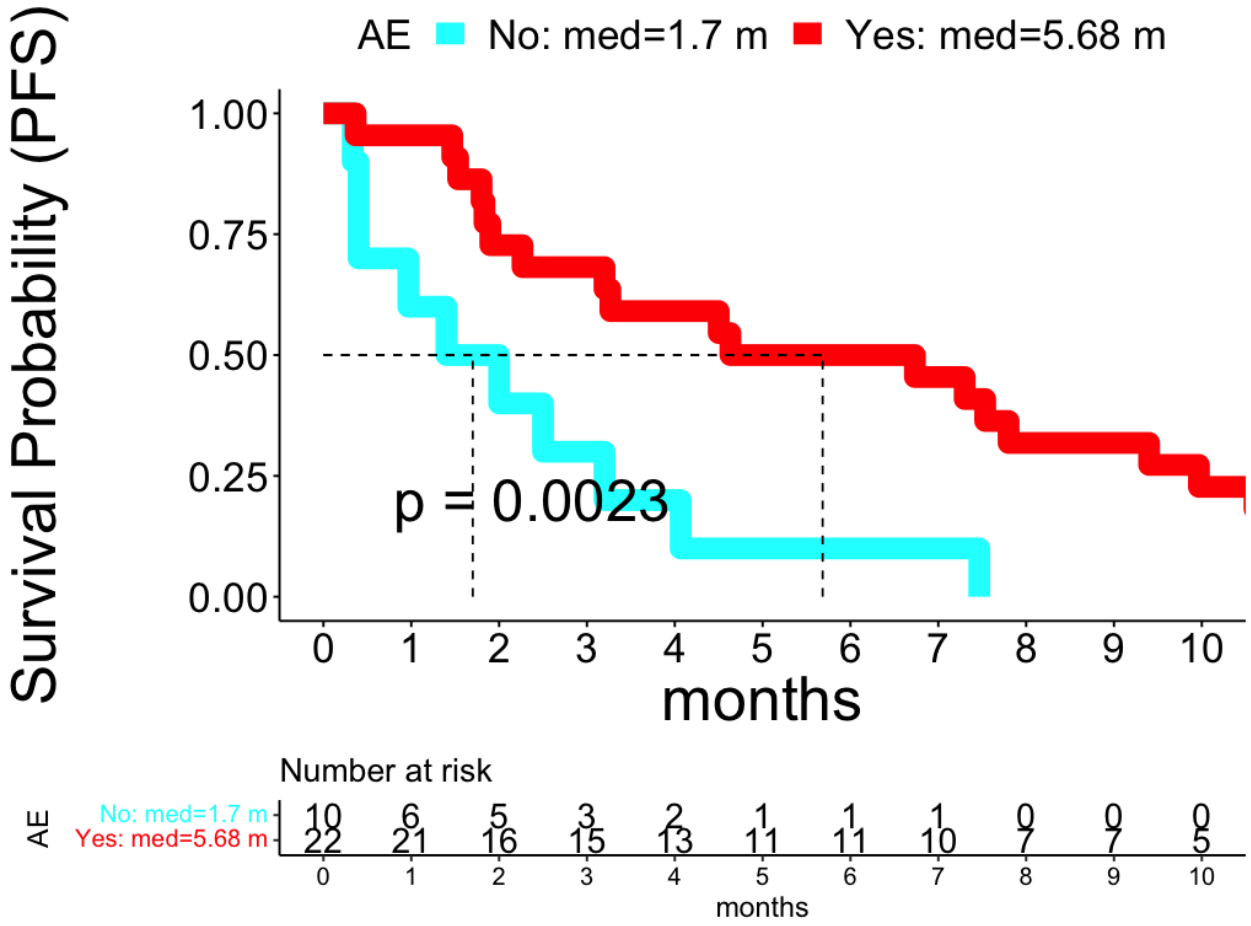
OS

Treatment related low-grade **overall early AE* biomarkers**



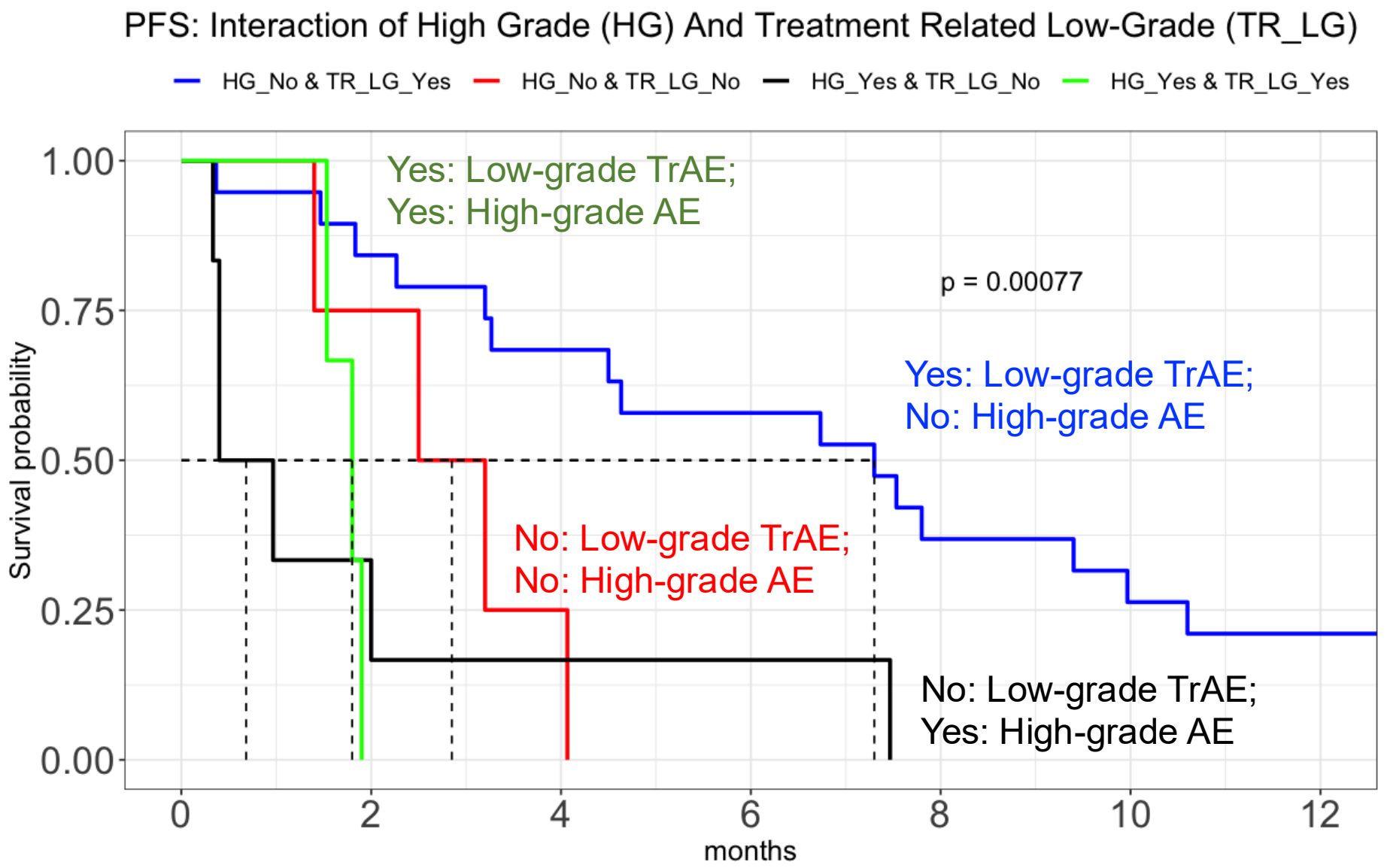
*Early AE: Occurrence of AE at or prior to day 30 of treatment

PFS
Treatment related low-grade **overall early AE* biomarkers**



*Early AE: Occurrence of AE at or prior to day 30 of treatment

Interaction of Treatment Related Low-Grade And High-Grade AEs



Strength

Analyzable and Informative AE predictive biomarker of clinical outcomes

AE-derived biomarkers					
		Measurement type			
		Occurrence	Sum of all unique AEs	Sum of all AEs	Sum of all AE duration
Grade type	Any grade	x	x	x	x
	Any treatment related grades	x	x	x	x
	Low-grade (1 or 2)	x	x	x	x
	Treatment related low-grade	x	x	x	x
	High-grade (3 or higher)	x	x	x	x
	Treatment related high-grade	x	x	x	x

Flexibility of Extension to Longitudinal AE Biomarkers

Potential opportunities of longitudinal AE biomarkers

- Biostatistical application opportunities
  - Mixed effect model to associate longitudinal AE data with demographic/clinical variables
  - Statistical joint model to utilize longitudinal AE data in analyzing/predicting survival data.
- Clinical application opportunities
  - Management of personalized toxicity profile by utilizing longitudinal AE data
  - Leverage of longitudinal AE data to predict clinical outcomes for early identification of disease and timely treatment intervention

Flexibility of Extension to Longitudinal AE Biomarkers (Cont.)

AE data at **one time point**

AE-derived biomarkers					
		Measurement type			
		Occurrence	Sum of all unique AEs	Sum of all AEs	Sum of all AE duration
Grade type	Any grade	x	x	x	x
	Any treatment related grades	x	x	x	x
	Low-grade (1 or 2)	x	x	x	x
	Treatment related low-grade	x	x	x	x
	High-grade (3 or higher)	x	x	x	x
	Treatment related high-grade	x	x	x	x

How??

Longitudinal AE data

Many Ways to Construct Longitudinal AE Data (A Challenge)

What type of longitudinal AE data is informative? (**need data mining**)

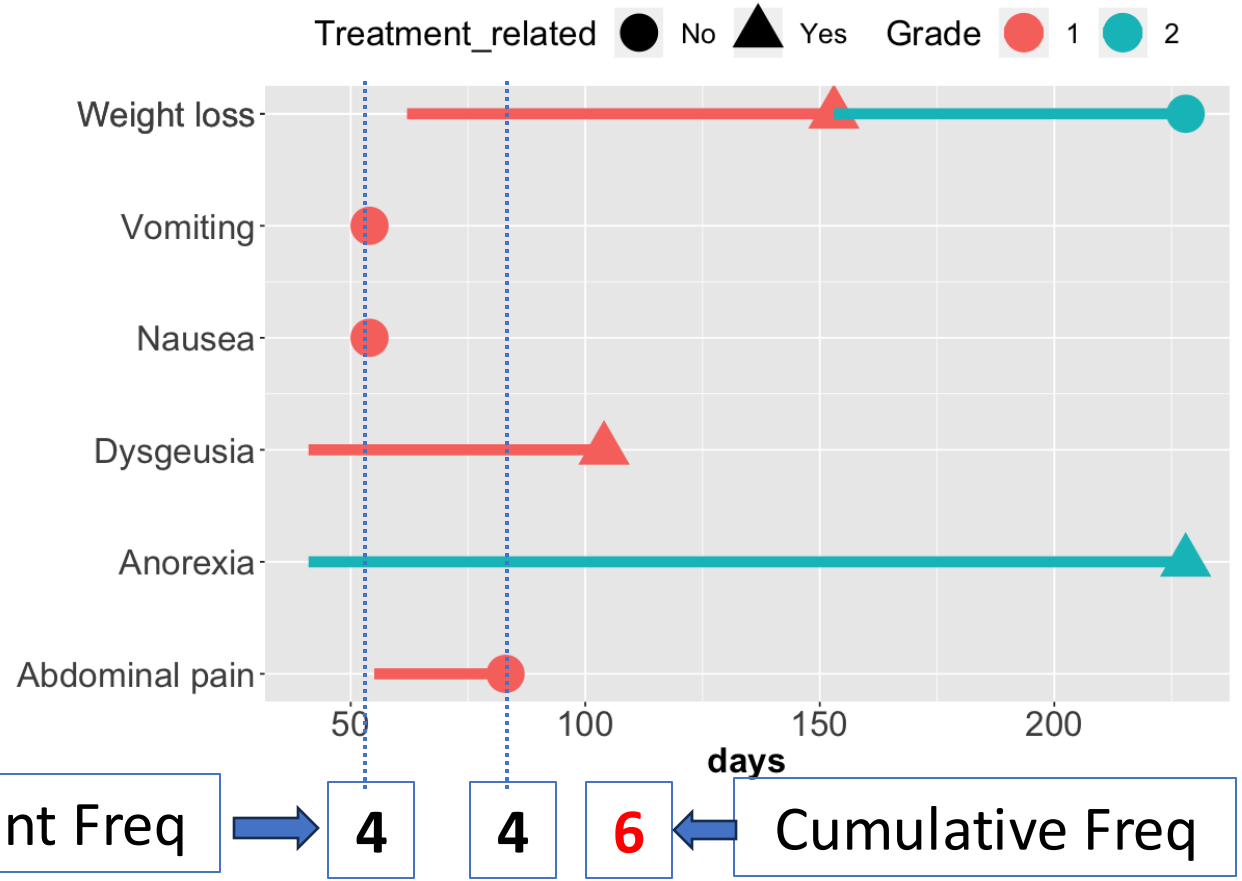
Examples:

Frequency:

- Timepoint
- Cumulative

Duration:

- Summation of all AE duration
- Duration of AE occurrence



One Statistical Application
Joint Model

- Dr. Rizopoulos's lecture note: https://www.drizopoulos.com/courses/EMC/ESP72.pdf
- Rizopoulos, D. (2012). Joint Models for Longitudinal and Time-to-Event Data, with Applications in R. Boca Raton: Chapman & Hall/CRC.
- R packages: JMbayes2

Joint Models

- Describe evolution of the covariate/marker over time for each patient.
- Then estimate evolutions in a Cox model

Joint Models (Cont.)

- Y_1 and Y_2 :Two outcomes of interest
- measured on a number of subjects for which
- Y_1 : Longitudinal data (continuous/binary/categorical variable)
- Y_2 : Survival data
- Goal: To jointly model both Y_1 and Y_2

Joint Models (Cont.)

Connection of joint model by random effect δ

$$p(y_{1,}y_2) = \int p(y_{1,}y_2|b)\, p(b) d\delta$$

$$= \int p(y_1|b)\, p(y_2|b) p(b) db$$

- Association between Y_1 and Y_2 by δ
- Y_1 and Y_2 conditional independence by δ

Joint Models: Our Application

Longitudinal Model:

$$y_i = \eta_i(t) + \varepsilon_i(t)$$

$$where\ \boldsymbol{\eta_i(t) = x_i(t)\beta + z_i(t)b_i}$$

Fixed effect Random effect

Survival Model:

$$h_i(t) = h_0(t)\exp(w_i(t)\gamma + \boldsymbol{\eta_i(t)\alpha})$$

AE Data Analysis by Joint Model

- Longitudinal data: Frequency of early low-grade treatment related adverse event (AE)

  Mixed effect model:

  fixed effect: AE= time + gender + age_at_on_study ← $x_i(t)\beta$

  random effect: random slope by subject ← $z_i(t)b_i$

  $\eta_i(t)$

  AE data (frequency): Poisson distribution

- Survival data: Progression-free survival (PFS)

  Cox proportional hazards model:

  PFS= gender + age_at_on_study + AE functional form ← $\eta_i(t)\alpha$

Results (Cox Model Only)

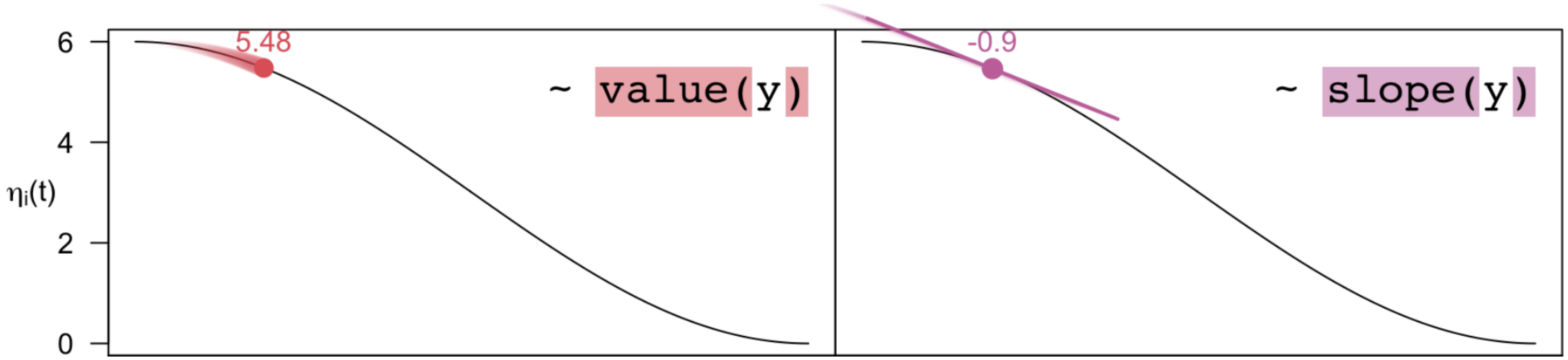
- Cox model only:

	coef	exp(coef)	se(coef)	z	Pr(>|z|)
AE Frequency	-0.796	0.451	0.326	-2.445	0.014
genderMale	-0.263	0.769	0.472	-0.557	0.578
age_at_on_study	-0.061	0.941	0.024	-2.572	0.010

Results (Joint Models)

Two types of AE function form (for AE frequency and cumulative frequency)

- Value
- Slope of AE frequency



Results (Joint Models)

AE Frequency

Survival Outcome: AE Frequency Value

	Mean	StDev	2.5%	97.5%	P
genderMale	-1.9309	2.3932	-7.2205	2.4589	0.3825
age_at_on_study	-0.0156	0.1052	-0.2103	0.2127	0.7913
value(freq)	-5.5319	3.6475	-12.5925	2.9948	0.1268

Survival Outcome: Slope of AE Frequency

	Mean	StDev	2.5%	97.5%	P
genderMale	-0.2176	1.4474	-3.1905	2.7454	0.8580
age_at_on_study	-0.1348	0.0726	-0.3065	-0.0232	0.0168
slope(freq)	-180.6852	102.9179	-430.8530	-37.4937	0.0104

Results (Joint Models)

AE Cumulative Frequency

Survival Outcome: AE Cumulative Frequency Value

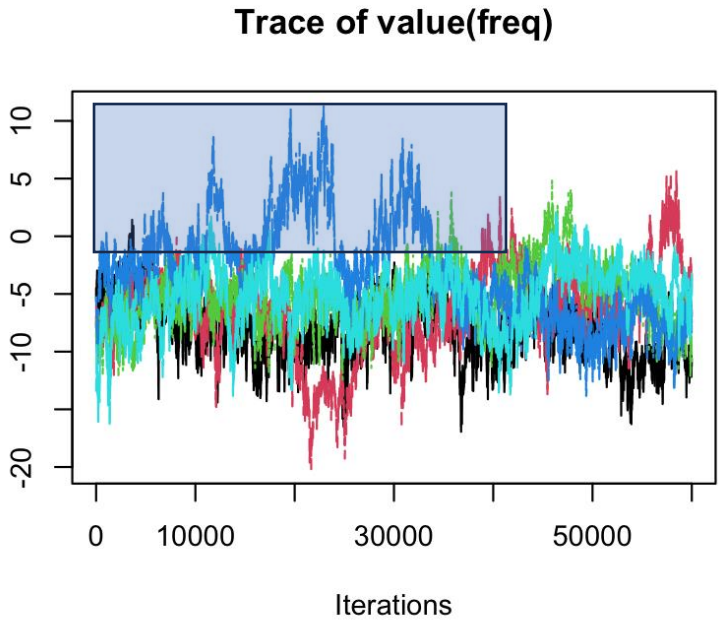
	Mean	StDev	2.5%	97.5%	P
genderMale	-1.9615	1.8596	-6.1373	1.2990	0.2444
age_at_on_study	-0.0337	0.0805	-0.1881	0.1408	0.5910
value(freq.cum)	-4.2138	1.9893	-8.6276	-1.0992	0.0020

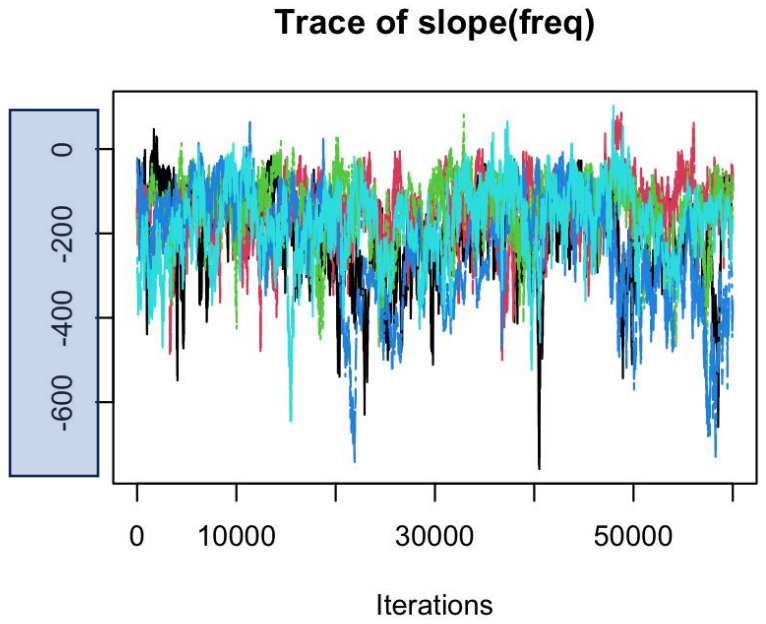
Survival Outcome: Slope of AE Cumulative Frequency

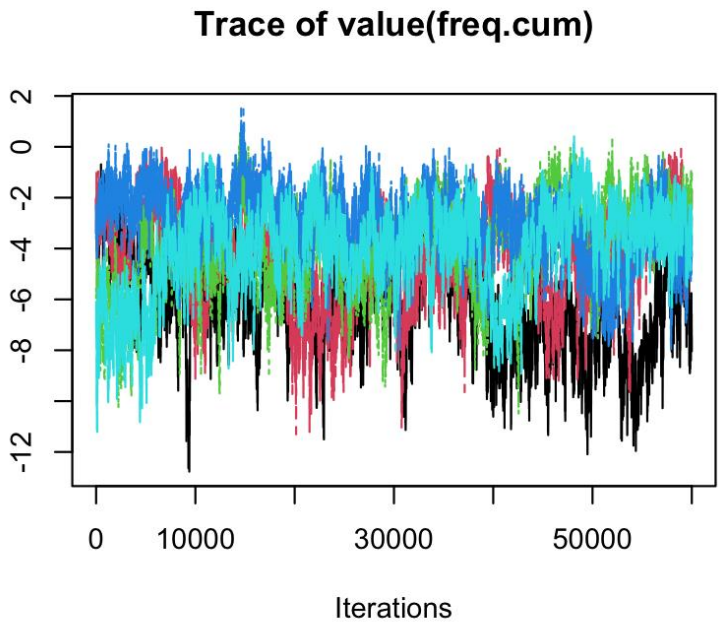
	Mean	StDev	2.5%	97.5%	P
genderMale	-0.2619	1.1330	-2.5507	2.0350	0.7958
age_at_on_study	-0.1202	0.0559	-0.2504	-0.0306	0.0074
slope(freq.cum)	-62.3537	34.2107	-148.4117	-15.0082	0.0011

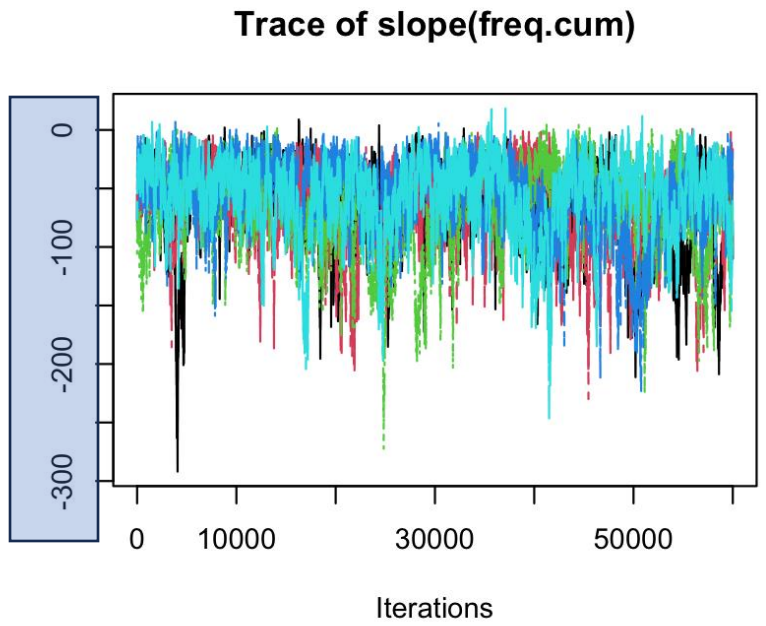
Diagnostic Plot

- Large variation in slope form
- More stable result in cumulative frequency









Bright Future of AE Data Mining

- Underutilized (only limited to descriptive toxicity report)
- Informative of clinical outcomes
- Great opportunity of data mining for data science community due to complexity of AE data

- Email: Dung-Tsa.Chen@moffitt.org

- Paper: PMID: 37173987
- R package: AdverseEvents (https://cran.r-project.org/web/packages/AdverseEvents/index.html)